

Norges miljø- og biovitenskapelige universitet
Fakultet for veterinærmedisin og biovitenskap
Institutt for kjemi, bioteknologi og matvitenskap

Masteroppgave 2014
30 stp

Stammespesifikk karakterisering av mikrobiota hos mødre og barn

Strain-Specific Characterization of the Microbiota in
Mothers and Children

Melissa Schanche

Forord

Denne masteroppgaven ble gjennomført ved Norges miljø- og biovitenskapelige universitet, ved instituttet for kjemi, bioteknologi og matvitenskap, med professor Knut Rudi og doktorgrad stipendiat Ekaterina Avershina som veiledere.

Først vil jeg takke professor Knut Rudi som aksepterte min forespørsel om å skrive min masteroppgave ved hans labgruppe for mikrobiell diversitet. Det har vært umulig å ikke bli smittet av din positive innstilling og entusiasme. Takk til medveileder Ekaterina Avershina for alt av teknisk hjelp på labben, og for god støtte under det skriftlige arbeidet. En stor takk til dere begge for den gode jobben dere gjorde med analysen av sekvenseringsdataene mine i MATLAB og veiledningen av det tekniske ved selve sekvenseringen. Dere var alltid tilgjengelige for å besvare «dumme» spørsmål, dere delte deres enorme kunnskap og viste stor tålmodighet. Jeg er rett og slett takknemlig for at dere hadde troen på meg. Til dere andre ved MiDiv gruppen; tusen takk for et fint og sosialt vårsemester, og for alle kaker og søtsaker dere delte med meg på mandagsmøtene våre.

Videre vil jeg takke Selma Kreso fra Genetic Analysis AS for hennes hjelp og veiledning ved innhenting av stammesamlingen brukt i denne oppgaven, «jentene» dine gjorde en god jobb.

Jeg vil også takke mine venner her ved IKBM, NMBU og hjemme i Drøbak; Cyril Frantzen for motiverende ord og vitenskapelige samtaler, som fort kunne spore av til en god latter. Silje-Marie Van Hoek og Camilla Kristoffersen for vennskapet vi har delt i fem studieår. En spesiell takk til deg Camilla, for at du er den du er! Jeg kunne ikke ha hat en bedre labpartner enn deg, og jeg er stolt over det vi har oppnådd sammen. Gjengen ved CrossFit Follo og Drøbak pole mafia, takk for at dere bidro til sosiale aktiviteter og energi når jeg trengte det mest.

Tusen takk til mamma og pappa for at dere alltid hadde troen på meg, og støttet meg uansett hva. Jeg hadde ikke vært der jeg er i dag uten dere. Sist men ikke minst, takk til min kjæreste Marius for at du alltid stilte opp, dine støttende ord hjalp meg til å holde ut i stressende øyeblikk.

Ås, mai 2014

Melissa Schanche

Sammendrag

En identifikasjon på art og stammenivå vil kunne gi utfyllende taksonomisk informasjonen utover den som allerede eksisterer basert på 16S rRNA genet. Denne informasjonen kan brukes til å gi et dypere innblikk i den bakterielle diversiteten i ulike mikrobiotaer hos mennesket. Spesielt vil dette gjelde i forhold til overføring av bakterier. Målet med denne masteroppgaven var derfor å kunne utvikle en stammespesifikk PCR som skulle gjøre det mulig å differensiere ulike bakteriestammer fra hverandre ved bruk av Illuminasekvensering. Denne metoden ble deretter brukt til å undersøke om bakteriell overføring mellom mor og barn kunne bekreftes. Vi evaluerte området mellom 16S-23S (ITS) rRNA genet og tmRNA genet som mulig kandidat for stammespesifikk identifisering. ITS primer paret ga et noe mer stabilt resultat ved amplifikasjon enn tmRNA primerparet, og ble valgt til videre evaluering av en stammespesifikk Illuminasekvensering. Evalueringen innbefattet test av et referansesett med tarmbakterier. Denne testingen viset at metoden hadde god sensitivitet og spesifisitet. For undersøkelse av bakteriell overføring fra mor til barn benyttet vi DNA i fra avføring og morsmelk, med opphav fra mor- barnpar som deltok i pro-PACT studiet. Parene som deltok i pro-PACT studie var delt inn i to ulike grupper, avhengig om mødrene hadde drukket Biola eller placebo melk fra svangerskaps uke 36 til tre måneder etter fødselen under amming. Morsmelk- og avføringsprøvene ble amplifisert med ITS primer paret, og amplikonene ble sekvensert med Illuminateknologien. Resultatene viste at ved å benytte det universelle ITS primerparet var det mulig å identifisere bakterier helt ned på stammenivå, og overføring mellom mor og barn kunne bekreftes. Sekvenseringen avdekket at 8 av de 10 mest dominante OTU'ene observert i morsmelk og avføringsprøvene analysert fra 19 mor- barnpar, var adressert til slekten *Streptococcus*. Den mest dominerende OTU'en totalt representerte en streptokokk som opprinnelig hadde blitt isolert fra munnhulen hos menneske. Disse streptokokkene ble funnet igjen i morsmelken til alle mødrene og var ofte, men ikke alltid, stabil over tid innenfor et individ.

Utviklingen av metoden befinner seg i en tidlig startfase. Denne studien viser at den innehar et potensiale som gjør det mulig å påvise bakterier på stammenivå, som kan utnyttes ved karakterisering av ulike bakterielle økosystemer.

Abstract

Identification at the species and strain level, could provide detailed taxonomic information beyond what already exist based on the 16S rRNA gene. This information can be valuable and provide a deeper insight into the bacterial diversity of different human microbiota. Especially will this apply in relation to transmission of bacteria. Hence, the goal of this thesis was to develop a strain specific PCR, which would make it possible to differentiate between strains by Illumina-sequencing. This method was then used to examine whether bacterial transmission between mother and child could be confirmed. We evaluated the area between the 16S and 23S (ITS) rRNA genes and the tmRNA gene as possible candidates for a strain-specific identification. The ITS primer pair gave a more stable result in terms of amplification, and was selected for further evaluation of a strain-specific Illumina sequencing. The evaluation included tests of a control set of intestinal bacteria. This testing showed that the method had good sensitivity and specificity. For examination of bacterial transmission from mother to child, we used DNA from feces and breast milk, originating from the mother and child pairs who participated in the pro-PACT study. The mother/infant pairs who participated in the pro-PACT study was divided into two groups, depending on whether the mothers had received Biola or placebo milk from 36 weeks of gestation, to 3 months postnatally during breastfeeding. Breast milk and feces samples were amplified with the ITS primer pair, and amplicons were sequenced. The results showed that, by using the ITS region as target, it was possible to identify bacteria down to strain level, and transmission between mother and child could be confirmed. The sequencing revealed that 8 out of the 10 most dominant OUT's in human milk and feces samples analyzed from 19 mother / infant pairs were addressed to the genus *Streptococcus*. Furthermore, the most dominant species turned out to be a *Streptococcus* originating from the human oral cavity. These streptococci were found to be present in breast milk samples recovered from all the mothers, and were often, but not always stable over time within an individual.

The method described still finds itself in an early stage of development. But, his study shows that it possesses a potential which allows detecting bacteria at the strain level, which can be utilized in characterizing various bacterial ecosystems.

Forkortelser

| | |
|-------------------|---|
| bp | Base par |
| dH ₂ O | Destillert vann |
| DNA | Deoksyribonukleinsyre |
| dNTP | Deoksynukleosid trifosfat |
| ds | Double trådet |
| ITS | Intern transkribert spacer region |
| NCBI | Nasjonalt senter for bioteknologisk informasjon |
| NMBU | Norges miljø- og biovitenskapelige universitet |
| PC | Prinsipal komponent |
| PCA | Prinsipal komponent analyse |
| PCR | Polymerase kjede reaksjon |
| pro-PACT | probiotika i PACT studiet |
| RNA | Ribonukleinsyre |
| rRNA | Ribosomal ribonukleinsyre |

Innholdsfortegnelse

| | |
|--|----|
| 1. Introduksjon | 1 |
| 1.1 Identifisering av bakteriearter | 1 |
| 1.2 Pro-PACT studiet | 3 |
| 1.3 Probiotika | 4 |
| 1.4 Human mikrobiota | 5 |
| 1.4.1 Bakteriell diversitet i morsmelk | 5 |
| 1.4.2 Kolonisering av spedbarnets tarmsystem | 6 |
| 1.5 Molekylære metoder for identifisering av mikrobiota | 8 |
| 1.5.1 Polymerase kjede reaksjon (PCR) | 8 |
| 1.5.2 Primerdesign | 8 |
| 1.5.3 DNA sekvensering med neste generasjons teknologi | 9 |
| 1.6 Multivariat analysemetode for kartlegging av mikrobiell diversitet | 10 |
| 1.7 Oppgavens målsetning og fremdriften i arbeidet | 12 |
| 2. Materialer og metoder | 14 |
| 2.1 Prøvematerialet | 14 |
| 2.2 Polymerase kjede reaksjon(PCR) | 15 |
| 2.2.1 Gradient PCR | 16 |
| 2.2.2 Opprensing av PCR-produkt med AMPure® XP | 16 |
| 2.3 DNA sekvensering | 18 |
| 2.4 DNA kvalifisering og kvantifisering | 18 |
| 2.4.1 Agarose gelelektroforese, kvalifisering av PCR produkt | 18 |
| 2.4.2 DNA kvantifisering med Qubit | 19 |
| 2.4 Dataanalyse | 19 |
| 3. Resultater | 21 |
| 3.1 Sammenligning av ITS og tmRNA primersettene | 21 |
| 3.2 Spesifisitet av ITS primere | 21 |
| 3.3 Illuminasekvensering | 24 |
| 3.3.1 Validering av metode | 24 |
| 3.3.2 Dominerende OTU'er i prøvematerialet | 26 |
| 4. Diskusjon | 30 |
| 4.1 Evaluering av Illumina ITS sekvenserings metode | 30 |
| 4.2 Påvisning og forekomst av dominerende OTU'er i pro-PACT prøvene | 31 |
| 4.2 Overføring fra mor til barn, og bakteriell opprinnelse i morsmelk | 33 |
| 5 Konklusjon og videre arbeid | 35 |

| | |
|---|----|
| Referanseliste | 36 |
| Vedlegg | 41 |
| Vedlegg A: Materialer og utstyr | 41 |
| Vedlegg B: Primer sekvenser | 42 |
| Vedlegg C: Bakteriestammer..... | 44 |
| Vedlegg D Agarosegel bilder | 46 |
| Vedlegg E Protokoller | 48 |
| Vedlegg F Utreknings eksempel for sammensetning av DNA bibliotek og antall genererte OTU'er i kontrollprøvene | 50 |
| Vedlegg G Forekomsten av de ti mest dominante OTU'ene i hvert mor- barnpar..... | 52 |

1. Introduksjon

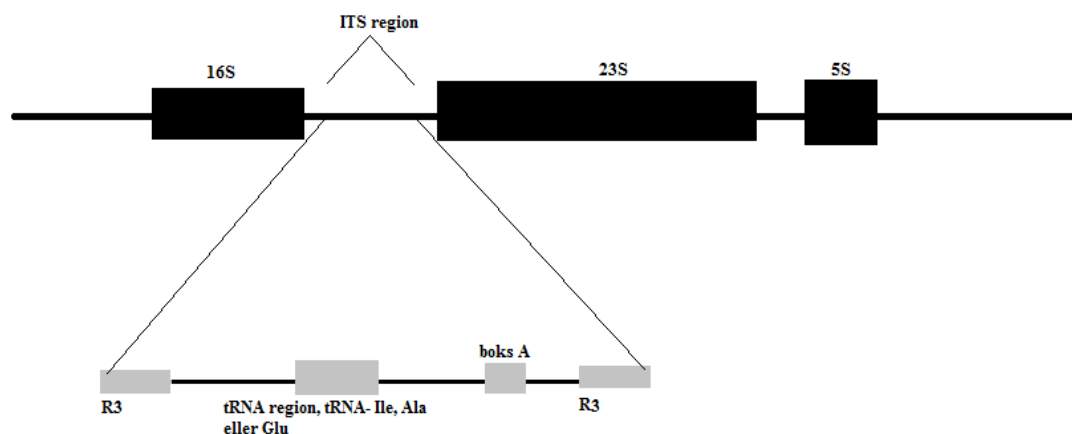
1.1 Identifisering av bakteriearter

Klassifiseringen av mikroorganismer har blitt basert på likhet i deres morfologi, levesett, metabolisme, evnen til å danne sporer og andre karakteristika som ikke baserte seg på kunnskapen om deres genotype. Bruken av stabile genomiske markører til bakteriell identifikasjon tok fart på 1980-tallet (Clarridge 2004), og den lille ribosomale 16S enheten har utmerket seg og regnes i dag som standard i å klassifisere bakterier (Vinje et al. 2014). Sekvensen for 16S rRNA genet er omtrent 1500 bp lang og består av konserverte og variable regioner. Dette genet har vist seg i å være nyttig i taksonomisk tildeling på grunn av dens innhold av informasjon, universale distribusjon hos prokaryote og dens konservative natur som ser ut til å være lite påvirket av evolusjonen (Lane et al. 1985). Sammenligning av 16S rRNA sekvenser gjør det mulig å skille bakterier på ulike taksonomiske nivåer. Dette oppnås ved bruken av universale primere som er komplementære til de konserverte områdene i genet. Til tross for dette kan det være vanskelig å identifisere to arter ved å bruke 16S rRNA genet alene, siden noen arter kan dele sekvenslikhet opp mot 99 prosent (Kim et al. 2014). For å overkomme dette problemet har andre områder i rRNA genet blitt utprøvd for å studere det fylogenetiske forholdet blant bakterier og deres identifisering. Saruta et al. (1995) brukte området mellom 16S og 23S rRNA genene for å identifisere ulike stammer av *Staphylococcus aureus*, mens Schönhuber et al. (2001) benyttet de konserverte sekvensene ved 5'- og 3'-endene av tmRNA genet til å designe universale primere for å identifisere gram-positive bakterier med lavt G+ C-innhold.

Sekvensen til det interne transkribert spacer (ITS) området som ligger mellom 16S-23S rRNA genene er mer variabel mellom bakterie arter enn selve 16S og 23S rRNA genene, noe som gjør den til en potensiell art og stammespesifikk sekvens (Saruta et al. 1995), en illustrasjon av ITS området vises i figur 1.1.

Det kan finnes seg flere ITS fragmenter innenfor ribosomal RNA operonet med varierende lengde, og antallet kan variere mellom 1 og 11 avhengig av art (Kostman et al. 1995). Variasjonen i lengde blir bestemt av tilstedeværelsen av flere funksjonelle enheter som tRNA gener og enheter som kan styre transkripsjonen (Gürtler & Stanisich 1996). I en studie som

omhandlet identifikasjon av melkesyrebakterier i meieriprodukter ble ITS sekvenser funnet med varierende lengde mellom forskjellige melkesyrebakterie-arter, og selv innenfor noen stammer av samme art. PCR amplifikasjonen av ITS fragmenter fra arter som inneholder flere kopier med variabel lengde, vil gi flere forskjellige produkter som resulterer i spesifikke mønstre av bånd ved en agarose gelelektroforese (Tilsala-Timisjärvi & Alatossava 1997). Det har blitt utviklet flere primere for bakteriell identifikasjon ved hjelp av ITS regionen som for *Mycobacterium*, *Streptococcus iniae*, *Escherichia coli*, *Salmonella* spp, *Bacillus* spp, deriblant også primere for deteksjon av *Bifidobacterium adolescentis* i yoghurt og avføring (Tsai et al. 2008).



Figur 1.1 Bakteriell ITS området innenfor rRNA operonet (svarte bokser), og organiseringen av dens funksjonelle regioner (grå bokser). De funksjonelle regionene viser tilstedeværelsen av tRNA gener som ikke er universell, og antall og type kan variere mellom arter. Ribonuklease 3 (R3) og boksA er med på å styre transkripsjon, heller ikke disse er universelle.

Zhang, J. et al. 2011 undersøkte også om sekvenser for tmRNA (tm) kan identifisere bakterier. Dette RNA molekylet innehar egenskapene til både tRNA og mRNA, derav navnet tmRNA. tmRNA har blitt funnet i en rekke bakterier og dets oppgave ser ut til å være involvert i translasjonen ved å tilgjengeliggjøre stoppkodoner for ribosomet dersom mRNAet mangler stoppkodon i 3'enden og videre fører til translasjonens avslutning (Zwieb et al. 1999). Kelley et al (2001) designet primere fra sekvenser funnet hos gamma proteobakterier og kunne ved hjelp av disse bestemme og identifisere 44 nye tmRNA sekvenser fra en rekke bakterielle samfunn.

1.2 Pro-PACT studiet

Prøvematerialet brukt i denne oppgaven stammer fra pro-PACT studien som er en understudie av PACT-studiet. The Prevention of Allergy Among Children in Trondheim (PACT) studie ble startet opp i år 2000 og inviterte alle gravide og foreldre til barn opp til 2 år i Trondheim til å delta i en ikke-randomisert, kontrollert, livsstils-intervensjonsstudie som ønsket å undersøke sammenhengen mellom primærforebyggende faktorer og forekomsten av allergi utvikling hos barn (Storro et al. 2010). Emnene angående livsstil for studiet var økt inntak av tran og fet fisk for kvinner under graviditeten og for spedbarn i løpet av deres to første livsår, reduksjon av foreldrenes røyking og redusert innendørs fuktighet. For å kunne innhente informasjonen om livsstilen vedrørende barna og deres foreldre, ble et spørreskjema fylt ut i løpet av svangerskapet, ved seks uker, ett år og to år etter fødsel.

Probiotika i PACT studiet (pro-PACT) er en randomisert, dobbelt blindet studie, det vil si at verken mødrene eller forskerne som gjennomførte studien visste om de fikk probiotika eller placebo. Studiet ble foretatt under samme tid som PACT studiet og de gravide kvinnene som deltok i pro-PACT studiet ble fordelt i to grupper, en probiotika- og en placebogruppe. Mødrene i studiegruppen fikk et probiotisk tilskudd i form av Biola, en syrnet -melk som inneholder de probiotiske bakteriene *Lactobacillus rhamnosus* GG (LGG), *Lactobacillus acidophilus* La-5 (La-5) og *Bifidobacterium animalis* subsp. *lactis* Bb-12 (Bb-12) (Dotterud et al. 2010), mens kontrollgruppen fikk utdelt placebo Biola i form av syrnet-melk uten probiotiske bakterier. Både den probiotiske- og placebomelken ble produsert slik at de skulle ha lik smak og konsistens, og ble utdelt i anonyme melkekartonger til mødrene i studiet. Studiet gikk ut på at kvinnene skulle drikke Biola fra svangerskapsuke 36 til tre måneder etter fødsel, barna skulle ammes under denne perioden.

Et spørreskjema som blant annet omfattet familie historie om atopi, kjønn, fødselsvekt, amming, prematur fødsel, alder til mor og eksponering for husdyr ble fylt ut av deltakerne under svangerskapet, ved seks uker, ett år og to år etter barnets fødsel. I tillegg ble det samlet inn prøver av avføring, blod fra navlestrengen og venøst blod fra barna etter 10 dager, tre måneder og ved ett- og toårs alderen, sammen med blant annet brystmelk fra mødrene. Når barna fylte to år, ble de undersøkt for eksem, astma og allergi for å se på effekten av probiotiske bakterier sammenlignet med placebo, med tanke på forekomsten av disse sykdommene mellom de to gruppene. I følge Dotterud et al (2010) viste resultatene at de

probiotiske bakteriene reduserte forekomsten av atopisk eksem hos barn frem til to års alderen med 40 %, men ingen signifikant effekt ble påvist med verken astma eller allergitest. Videre kan det nevnes at 21 % av barna i probiotikagruppen og 34 % av barna i placebogruppen fikk diagnosen barneeksem frem til to års alderen. Der barna i probiotikagruppen som likevel fikk eksem var eksemet mindre alvorlig og mild i følge Nothingham Eczema Severity Score (Dotterud et al. 2010).

1.3 Probiotika

Probiotika defineres som levende mikroorganismer som kan gi helsefremmende effekter på verten utover vanlig ernæring ved en viss dosering (Goldin 1998) . Probiotika går også under navnet melkesyrebakterier, og disse kan benyttes i matvarer eller inntas gjennom kosttilskudd. For å kunne feste seg til tarmcellene eller mukus må bakteriene inneha egenskaper som gjør det mulig å overleve passasjen gjennom magesekken og tolvfingertarmen for å komme frem til tarmen uten å bli ødelagt av syrebarrieren. Midlertidig kolonisering eller addisjon av de probiotiske bakteriene til mukosa gjør det mulig for dem å utøve sin virkning. Melkesyrebakteriene blir ikke alltid gjenkjent av immunsystemet i tarmen som kroppsegne, og heftingen til mukosa vil ikke bli godtatt (Birkeland & Tandberg 2003). I tillegg kjemper de om plassen med bakterier som allerede er kolonisert i tarmen ved blant annet å skille ut ugunstige stoffer mot andre mikroorganismer som bakttereosiner, organiske syrer og hydrogenperoksid (Shah 2001). Bakterien *Lactobacillus acidophilus* La-5 skiller ut laktat, eddiksyre og hydrogenperoksid ved glukosefermentering som fører til at pH i tarmen synker og dermed gjør miljøet mindre gunstig for vekst av patogene (Lee & Salminen 2009). Ved regelmessig inntak av utvalgte melkesyrebakterier kan disse bidra til å endre eller normalisere sammensetningen av tarmfloraen, som kan føre til at omsetningen av næringsstoffer og andre bioaktive stoffer endres (Collins & Gibson 1999; Goldin 1998). Bedre egenskaper i kamp om plassen på tarmveggen og i kampen om næring er andre faktorer som bidrar til den hemmende effekten på mikroorganismer i tarmen som ikke er gunstige for helsen. De vanligste probiotiske stammene tilhører laktobasiller eller bifidobakterier som finnes naturlig i menneskets tarmflora. Disse blir brukt til å syrne probiotiske produkter som Biola og Cultura. *Lactobacillus rhamnosus* GG eller LGG, *L. casei* Shirota, *L. johnsonii* La1, *L. plantarum* 299v og *Bifidobacterium* Bb-12 er de mest brukte og studerte stammene av probiotika (Ouwehand et al. 2002). Bakteriene blir tilsatt i store mengder til produktene for å øke sjansen

av midletidig kolonisering, hvor antallet varierer mellom 10 til 100 millioner bakterier per ml (Birkeland & Tandberg 2003). Innvirkning på tarmens immunsystem, toleranse for melkesukker, lindring av lettere fordøyelsesproblemer, forebygging mot tykktarmskreft, atopisk eksem og diaré er noen av de helserelaterte effektene studier prøver å belyse (Aso et al. 1994; Bennett et al. 1996; Kalliomäki et al. 2001; Vanderhoof et al. 1999). Av disse effektene er det spesielt virkningen mot ulike diarètilstander og forebygging av fordøyelsesforstyrrelser som ved inntak av antibiotika, rotavirus, vandig diaré, *Clostridium difficile* infeksjoner, *Shigella* og *Salmonella* hos barn og *E. coli* ved reisediarè som har stått i fokus. Probiotiske bakterier ser ut til å styrke den naturlige tarmfloraen og hindrer patogener i å trenge igjennom epitellaget ved å normalisere tarmpermeabiliteten.

For denne oppgaven er den immunregulerende effekten til probiotika sentral, da pro-PACT studiet omhandler effekten av probiotisk melk og allergi utvikling. Litteraturen angir at store deler av det humane immunapparatet befinner seg i eller har tilknytning til tarmen, og at probiotiske bakterier som fester seg til tarmen kan stimulere gastrointestinalt assosiert lymfoid vev på en positiv måte (Andersson et al. 2001). Ved å påvirke den mikrobielle balansen kan probiotiske bakterier fremme antiallergene prosesser og stimulere produksjonen av signalstoffet cytokin-Th1 og serum IgA i tarmen, samt generering av transformerende vekstfaktor β (Rautava et al. 2002). Konfrontasjonen mellom mikrober og deres antigener i mage-tarm-kanalen begynner umiddelbart etter fødselen, derfor utgjør kommensale mikrober den tidligste og største stimulans for utvikling av gastrointestinalt-assosiert lymfoid vev (Isolauri et al. 2001).

1.4 Human mikrobiota

1.4.1 Bakteriell diversitet i morsmelk

Morsmelk er barnets første møte med næring etter fødselen. I tillegg til å gi barnet essensielle næringsstoffer viser litteraturen at morsmelk kan være med å beskytte det mot infeksjoner (Morrow & Rangel 2004). Denne beskyttende egenskapen kan komme av ulike stoffer i melken som antistoffer, immunokompetente celler, oligosakkarider, lysosomer, laktoferin, samt andre glykoproteiner, peptider og fettsyrer som virker antimikrobielt (Martín et al. 2007). En annen kilde til denne effekten kan være morsmelk som en kilde til kommensale og

probiotiske bakterier som kan kolonisere spedbarnets tarm og inaktivere patogene mikroorganismer (Lara-Villoslada et al. 2007).

Studier på morsmelk viser at sammensetningen av bakterier er heterogen og spesifikk for hver enkelt mor (Collado et al. 2009). Stafylokokker, streptokokker, mikrokokker, laktobasiller, enterokokker, bifidobakterier og slekten *Bacteroides* er blant de mest dominerende bakterieartene funnet i morsmelk. Melkesyrebakterier er gunstige da de innehar egenskaper til å motvirke veksten av et bredt utvalg med patogene bakterier jamfør seksjon 1.3 om probiotika. Det er også funnet *E.coli* og *Clostridium* som er blant de første artene til å kolonisere spedbarnets tarm (Martín et al. 2007). Tilstedeværelsen av disse bakteriene i brystmelk er i tråd med utsagnet om at spedbarnets tarm-mikrobiota gjenspeiler den i brystmelk, og disse kan bli overført gjennom amming. Denne påstanden støttes også av Jimenez et al (2008) som har sett at de første artene som koloniserer spedbarnets tarm er blant de mest dominante i morsmelk. Dette indikerer en sammenheng mellom amming og tilstedeværelsen av enkelte bakterie arter i barnets tarmmikrobiota.

Det finnes flere teorier om hvordan brystmelk blir kolonisert. En forklaring er at enkelte arter stammer fra tarmen til mor og flytter seg til melkekjertlene via tarm epitelet, en prosess som er med på å stimulere immunforsvaret positivt (Albesharat et al. 2011). Martín et al (2007) observerte *Pseudomonas synxantha* i brystmelk. Tilstedeværelsen av *P. synxantha* har åpnet opp for teorien om at spedbarn mottar bakterier gjennom fødselen som overføres videre til morsmelk gjennom amming, da denne bakterien dominerer kjeden hos friske kvinner.

1.4.2 Kolonisering av spedbarnets tarmsystem.

Menneskets mage-tarm-kanal er kolonisert av ulike mikroorganismer fra munnhulen til endetarmen. Litteraturen gir et ulikt bilde av mangfoldet som mikrobiotaen omfatter, en diversitet på 400-500, 800 og 1000 ulike arter er nevnt (Bäckhed et al. 2005; Noverr & Huffnagle 2004; Rajilić-Stojanović et al. 2007). Dette økosystemet inneholder en høyere tetthet av celler enn noen annet kjent mikrobielt økosystem, og en unik sammensetning er funnet i hvert enkelt individ (Albesharat et al. 2011). Bakteriene i tarmen bidrar med å bryte ned fiberholdig mat, fettsyrer og syntetiserer vitaminer og mineraler, deriblant vitamin K. Barnets tarmsystem er sterilt så lenge det befinner seg i livmoren, men rett etter fødsel starter en bakteriell koloniseringsprosess som varer frem til 2-4 års alderen (Biasucci et al. 2010).

Men det er funnet bakterier i fostervann og i blod fra navlestrengen som indikerer at disse har kolonisert spedbarnets hud og fordøyelses system før fødselen (Martín et al. 2007). Mikrobiotaen endrer seg stadig i denne perioden, helt til den stabiliserer seg og ligner på sammensetningen funnet hos voksne. Bakteriene stammer primært fra fødselskanalen, morsmelk og omgivelsene spedbarnet befinner seg i. Ulike faktorer påvirker mikrobiotaens sammensetning og kan deles inn i vertspesifikke faktorer som pH, enzymer og mukosaens oppbygning. Ikke vertspesifikke faktorer som kosthold, bruk av medisiner og miljø, samt bakterielle faktorer som addisjons evne, enzymer og metabolsk kapasitet er også med å bestemme den bakterielle komposisjonen (Penders et al. 2013). Spedbarnets tarmsystem blir først kolonisert av aerobe og fakultativ anaerobe bakterier som inkluderer enterobakterier, enterokokker og stafylokokker. Når miljøet i tarmen senker oksygennivået gir det grunnlag for spredning av anaerobe bakterier som bifidobakterier, clostridia og bacteroider (Abrahamsson et al. 2012). Tarmens mikrobiota påvirker utviklingen til immunsystemet hos nyfødte barn, både det medfødte og det adaptive systemet. Den er med på å utvikle og opprettholde immuntoleransen som bidrar til å forebygge allergi og autoimmune sykdommer. Litteraturen viser at den bakterielle sammensetningen hos barn med allergi utvikling skiller seg fra friske barn i at diversiteten er lavere og består av et færre antall laktobasiller, bacteroider, *E.coli* og bifidobakterier, mens koloniseringen består av et høyere antall *Firmicutes*, *Staphylococcus aureus* og *Clostridium difficile* (McLoughlin & Mills 2011; Watanabe et al. 2003). Tilstedeværelsen av streptokokker er som nevnt en naturlig del av mikrobiotaen hos friske spedbarn, men er sjeldent eller aldri funnet hos barn som har allergi (Collado et al. 2009). Fødselsmåten og om barnet blir ammet, samt påvirkningen av disse faktorene på barnets mikrobiota og sykdomsutvikling har fått mye oppmerksomhet. Barn født med keisersnitt har høyere sjanse for å utvikle allergi. Det er fordi de ikke kommer i kontakt med morens vaginale og fekale bakterieflora og deres primære kolonisering består av bakterier som stammer fra sykehuset og ikke-maternal hud, noe som senker diversiteten hos spedbarnets tarmflora og forsinker koloniseringen av gunstige tarmbakterier. Dette er i tråd med *C.difficile* funnene hos barn med allergi, da dette er en typisk sykehusbakterie (Penders et al. 2013). Som nevnt over er morsmelk en kilde til kommensale og probiotiske bakterier som er med på å avgjøre sammensetning av barnets mikrobiota ved å øke diversiteten og dermed stimulere den immunologiske toleransen som er med på å gjøre barnet bedre rustet mot eksponering av allergener og ugunstige mikrober under dets utvikling.

1.5 Molekylære metoder for identifisering av mikrobiota

1.5.1 Polymerase kjede reaksjon (PCR)

PCR er en *in vitro* metode for å syntetisere spesifikke DNA sekvenser ved hjelp av enzymet polymerase, de fire deoxynukleotidene som er byggesteinene i DNA, og to korte oligonukleotid primere. Primerne er spesifikke og hybridiserer til komplementære sekvenser i templat DNA. Den voksende tråden syntetiseres ut ifra primerens 3'-OH ende og er komplementær til templat tråden. Sekvensen som kopieres er begrenset på hver side av plasseringen til primerne. Fordi produktene som har blitt syntetisert ved en syklus fungerer som templat i neste, vil antall kopier for målsekvensen dobbles ved hver syklus (Mullis et al. 1992). Fasene i reaksjonen utføres ved ulike temperaturer og hver syklus består av en denatureringsfase av templatet til enkeltrådet DNA, primer hybridisering og en ekstensjonsfase ved polymerasens temperatur optimum, som fører til en eksponentiell akkumulering av PCR-produkt. Som templat er det nok med ett eneste DNA-molekyl som inneholder den sekvensen man ønsker å amplifisere (Erlich 1989).

1.5.2 Primerdesign

Utformingen av primere har innvirkning på PCR med hensyn til spesifisitet og utbyttet av det amplifiserte produktet. Et problem som forekommer er primere som binder til ukorrekte regioner av DNA, de oppfører seg uspesifikke og kan gi uventede produkter. Størrelsen på primere som overgår 16-28 base par, som anses å være optimal lengde for primere, bidrar til at primerne binder dårligere til templatet i tillegg til at sjansen for uspesifikk primer binding kan øke (Chuang et al. 2013). For å unngå dette er det mulig å utføre en to-steps PCR prosedyre som innebærer å bruke produktet fra den primære PCR som templat for en sekundær PCR. Denne måten å sette opp reaksjonene på senker sjansen for at PCR-produktet inneholder kontaminering av uønskede produkter som primer dimer, hårnål strukturer og alternative sekvenser. Hybridiseringstemperaturen til primerne er et kritisk punkt i å unngå dannelsen av ikke-spesifikke sekundære strukturer. Optimalisering av temperaturen kan bestemmes ved hjelp av en gradient-PCR. Blokken til gradient-PCR maskinen gir en gradient-funksjon som i en enkelt kjøring evaluerer opp til 12 forskjellige hybridiseringstemperaturer.

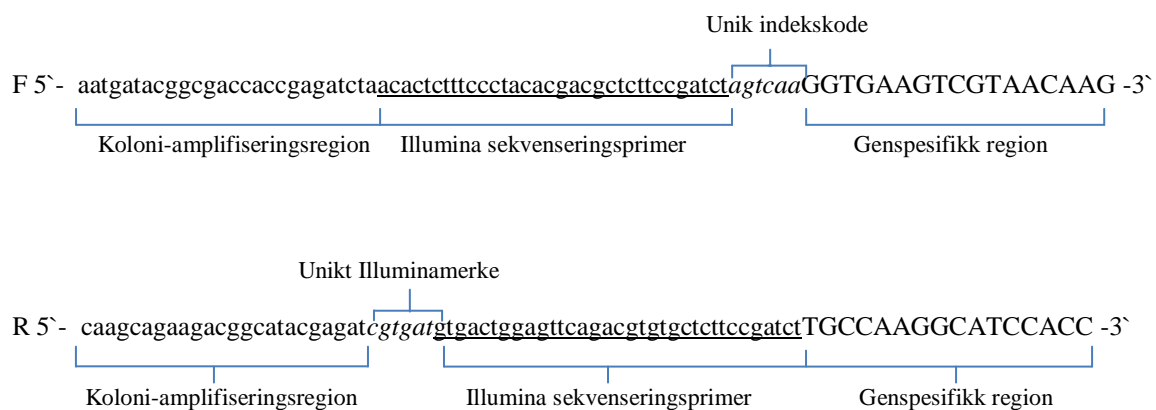
Andre kriterier som må tas hensyn til ved primernes karakteristikk er lengde differansen mellom primerne, deres GC innhold og nukleotide sammensetningen i 3'-enden (Dieffenbach et al. 1993).

1.5.3 DNA sekvensering med neste generasjons teknologi.

DNA sekvensering betegnes som en detaljert beskrivelse av rekkefølgen av de kjemiske byggeklossene eller baser i DNA. Bestemmelse av taksonomien til organismer og sammenligning av disse, kartlegging og analyse av ny gener, og sporing av mutasjoner er noen av anvendelsesområdene. Den automatiserte Sanger metoden regnes som "førstegenerasjons"-teknologi, og nyere metoder omtales som neste generasjons sekvensering (NGS) (Metzker 2010). NGS kjennetegnes ved dens evnen til å produsere store mengder av data ved lave kostnader på kort tid. Den store mengden data produsert av NGS representerer også en betydelig utfordring for datalagring, bioinformatisk verktøy og administrasjonsløsninger (Zhang et al. 2011). Det første NGS utstyr ble tilgjengelig i 2004 og i dag er det flere plattformer tilgjengelig på markedet for dybdesekvensering, som Roche/454, Illumina/Solexa, Life/APG og Helicos BioSciences (Morey et al. 2013).

I denne oppgaven ble sekvensering av pro-PACT prøvene utført på Illumina MiSeq plattformen. Plattformen benytter seg av sekvensering ved syntese metoden som blir utført ved en bro-PCR amplifikasjon prosess. Bro-amplifisering tillater amplifisering av et stort antall DNA-fragmentene samtidig, ved hjelp av en flytcelle med en fast overflate som er belagt med oligonukleotider. Oligonukleotidene er komplementære til adapterne tilsatt til DNA-prøvene under bibliotekkonstruksjonen. Ved å feste en spesifikk DNA-sekvens i form av en barkode til hvert fragment gjør at de kan multiplekseres. Barkodene kan tilsettes under konstruksjonen av biblioteket ved å feste dem til adaptere, og disse barkodeadapterne kan adderes til DNA-fragmentene ved en amplifisering med primere som inneholder barkodeadapterne. En illustrasjon av primere for Illuminasekvensering vises i figur 1.2. Sekvensene til barkodene er derfor spesifikke for ulike prøver som gjør det mulig for plattformen å skille dem fra hverandre (Shendure & Ji 2008). DNA-fragmentene fra biblioteket blir denaturerte, og så festet til flytcellens overflate. Oligonukleotidene er bundet til overflaten av flytcellen med sin 5'-ende, slik at 3'-enden er tilgjengelig for polymerasen.

Dette resulterer i dobbelt-trådet DNA (dsDNA) som er kovalent bundet til overflaten som igjen blir denaturert, og det enkelt trådet DNA kan vippest over til nærliggende primere og hybridisere med disse for å danne en bro. Ekstensjonen av ampliconene oppnås gjennom en polymerase og resulterer i dobbelt-trådig broer med en bestemt posisjon på flytcellen. Disse broene denatureres til to kopier av kovalent bundet enkelt-trådet templat. Som nevnt finner sekvenseringen sted på overflaten av flytcellen og polymerasen benytter reversibel-terminerende baser med fluoriserende egenskaper (Ansorge 2009). For å oppnå en naturlig konkurranse mellom basene som bidrar til korrekt inkorporering, tilsettes basene samtidig. Kun en base vil bli inkorporert ved hver syklus, derfor betegnes denne metoden som singel nukleotide addisjon (SNA). Etter basens inkorporering registreres det fluoriserende signalet som korresponderer til basen identitet ved fotografering. Kolonidannelsen forsterker signalet slik at det er mulig for kameraet å fange det opp. Etter deteksjonen blir 3' blokkingen kjemisk fjernet og prosessen kan starte på nytt (Mardis 2008).



Figur 1.2 En illustrasjon av Illumina Forward og Illumina Reverse primerpar. Primerne er designet slik at den genspesifikke regionen tilsvarer primersekvensene til det genet som skal amplifiseres. Bindingen til flytcellen på Illumina plattformen opprettes ved hjelp av koloniamplifiseringsregionen som er komplementær til sekvenser på flytcellen. Reverseprimeren inneholder et unikt Illumina merke og hver forward primer inneholder et unikt primer merke for at separasjon og gjenkjenning av koloniene skal være mulig.

1.6 Multivariat analysemetode for kartlegging av mikrobiell diversitet

Analyse av multivariat data er en stor del av dataanalyse. Multivariat data består av mange forskjellige attributter eller variabler registrert for hver observasjon. For å behandle sekvenseringsdataen i denne oppgaven ble Prinsipal komponent analyse (PCA) benyttet.

Den sentrale ideen bak PCA er å redusere dimensjonaliteten i et datasett som består av et stort antall variable som er forbundet med hverandre (Wood 2009). Samtidig ønsker man å beholde så mye som mulig av variasjonen tilstede i datasettet. Dette oppnås ved å omforme et nytt sett med variabler, disse betegnes som prinsipal komponenter (PC) og er ukorrelerte og ordnet på en slik måte at de første prinsipal komponentene beholder mesteparten av variasjonen som er tilstedet i de opprinnelige variablene. Prinsipal komponenter blir estimert fra en korrelasjonsmatrise eller egenvektorene til kovariansen av de originale variablene. Det er vanlig å bruke korrelasjonsmatrisen når variablene i datasettet har ulike enheter eller om ulike variabler har ulike varians (Jolliffe 2005). Korrelasjonsmatrisens egenvektorer tilfører vekten som skal til for å beregne PC'ene mens egenverdiene måler mengden av variasjonen gitt ved hver PC, og vil være størst for de første PC'ene og mindre for de påfølgende PC'er. En egenverdi større enn en indikerer at PC-ene står for mer varians enn det en av de opprinnelige variablene i original dataen stod for (Ku et al. 1995). Dette kan brukes til å bestemme hvilke PC'er som skal beholdes videre.

Antall prinsipal komponenter blir bestemt på en slik måte at mesteparten av variasjonen i datasettet beholdes med færrest mulig komponenter. Komponentene ordnes etter deres avtagende betydning slik at den første PC utgjør så mye av variasjon som mulig, og hver påfølgende komponent står for litt mindre variasjon. For å kunne reprodusere den totale variasjonen i systemet av de opprinnelige variablene blir alle PC'ene tatt i betraktning, men hvis de første PC'ene står for 80-90 % av variasjonen er målet om dimensjon reduksjonen oppnådd (Abdi & Williams 2010). Det er mulig å lage en Scree Test hvor egenverdiene blir plottet mot den korresponderende PC'en. Dette plottet illustrerer endring i hastighet for størrelsesorden til egenverdiene for PC'ene. Punktet hvor kurven bøyes indikerer det maksimale antallet PC som er aktuelt å bruke i analysen.

PC «loadings» er korrelasjonen mellom en PC og den originale variabelen, og estimerer informasjonen de deler seg imellom. Dette gir informasjon om betydningen av hver variabel med tanke på bidrag til variasjon til PC'en. Viser en variabel høy korrelasjon med for eksempel PC1 kan det indikere at variabelen assosieres med retningen til den maksimale mengden variasjon i datasettet. Mer enn én variabel kan ha høy korrelasjon med PC1. En sterk korrelasjon mellom en variabel og PC2 viser at den variabelen er ansvarlig for den nest største variasjonen i datasettet loddrett på PC1, og så videre. Hvis en variabel ikke korrelerer til noen av PC-ene, eller bare korrelerer med den siste PC tyder dette vanligvis på at den variabelen har liten eller ingen innflytelse på variasjonen i datasettet. Derfor kan PCA indikere hvilke

variabler i et datasett som er viktige og hvilke som kan være av liten betydning (Moore 1981; Smith 2002). PC-score er utledet av den sammensatte score beregnet for hver enkelt observasjon basert på egenvektoren for hver PC. Relasjonen mellom PC-score og PCA-loadinger for to PC-er kan bli illustrert i et Bi-plot. Dette plottet visualiserer relasjonene mellom observasjoner og variabler i multivariat data (Graffelman & van Eeuwijk 2005). For å lage et bi-plot bør dataene bli sett på som en matrise hvor kolonner representerer variable mens rader representerer observasjoner, og disse to dimensjonene plottes samtidig.

1.7 Oppgavens målsetning og fremdriften i arbeidet

16S rRNA genet gir ikke god nok oppløsning på stammenivå. Identifikasjon av stammer er viktig med tanke på kartlegging av bakterieoverføring fra mor til barn. Hovedmålsetningen med denne oppgaven var å evaluere primere som gjør det mulig å utvikle en stammespesifikk PCR som kunne brukes til identifikasjon av bakterier i morsmelk og avføring ved hjelp av sekvensering med Illumina teknologien.

Spesifikke mål for å oppnå hovedmålet var:

- Å evaluere utvalgte primersett med tanke på deres spesifisitet som kunne brukes til å utvikle en stammespesifikk Illumina-sekvenserings-metode.
- Å undersøke overføringen av bakteriearter fra mor til barn ved hjelp av prøver fra pro-PACT studiet.

Fremgangen i arbeidet illustreres i figur 1.3.

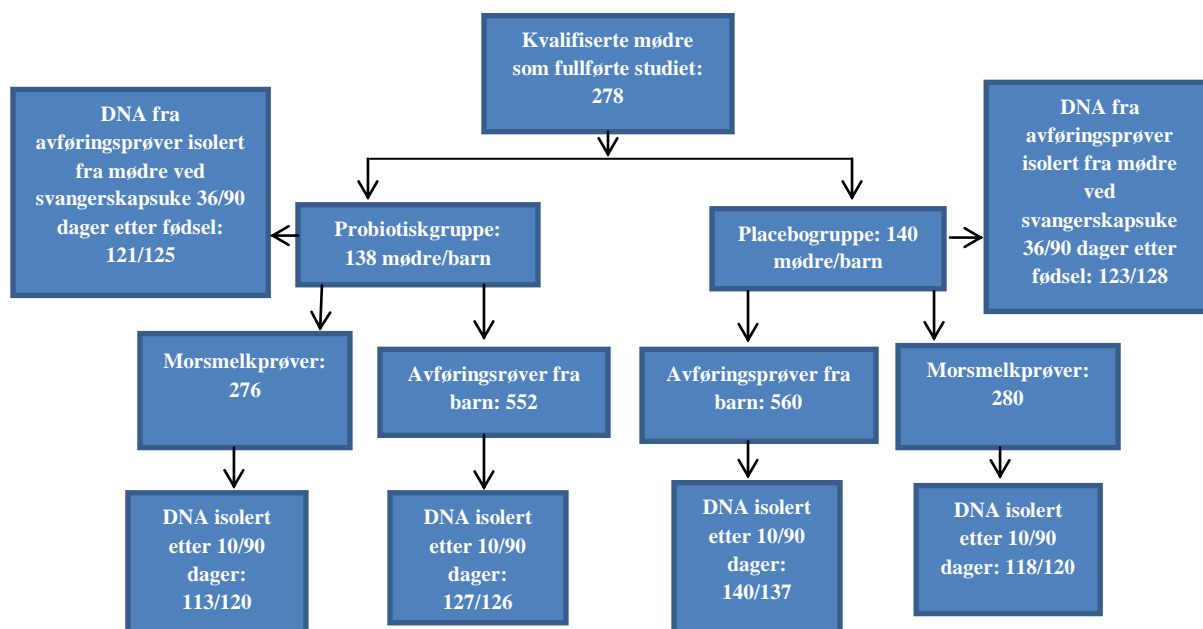


Figur 1.3 Flytskjema for denne oppgaven. Viser en skjematisk fremstilling av hovedgangen og delforsøk i arbeidet. Del a) illustrerer forarbeider som ble gjort før forsøkene utført med pro-PACT prøvene. Pilene indikerer forløpet av delforsøkene gjort i denne delen. Valg av primerpar og PCR optimalisering ble gjort i forarbeidet. Del b) illustrer forsøkene utført med pro-PACT prøvene og DNA fra Genetic Analysis AS etter evalueringen av hvilket primerparet som egnet seg for en stammespesifikk PCR. Pilene viser hvilke forsøk som har blitt gjort med templatet i boksen pilen tar utgangspunkt i. Runder bokser indikerer at forsøket beskrevet kun ble utført med det prøvematerialet beskrevet i boksen der pilen peker ut ifra.

2. Materialer og metoder

2.1 Prøvematerialet

Prøvematerialet brukt i denne oppgaven stammer fra pro-PACT studiet og Genetic Analysis AS. Avføring- og morsmelk DNA fra pro-PACT var blitt isolert ved bruk av LGC MagTM DNA ekstraksjonskit (LGC Genomics, UK) i henhold til instruksjoner fra produsenten. Av de 415 mødrene som var kvalifisert til studien var det 278 mødre som fullførte. Avføring og morsmelk ble isolert fra 138 mødre/barn par tilhørende gruppen som hadde fått Biola, og fra 140 mødre/barn par som tilhørte placebogruppen. Det ble avgitt morsmelksprøver to ganger fra de 278 mødrene, en prøve 10 dager etter fødsel, og en prøve 90 dager etter fødsel. Fire avføringsprøver ble avgitt fra det respektive barnet ved 10 dager, 90 dager, ett år og to år etter barnets fødsel. Det ble i tillegg avgitt til sammen to avføringsprøver fra mødrene, en ved svangerskapsuke 36 og en 90 dager etter fødsel. Fra disse prøvene ble det isolert DNA fra 1516 prøver totalt. Mødre/barn parene ble valgt tilfeldig ut ifra DNA prøvene som var tilgjengelige, 10 par fra probiotikagruppen og 10 par fra placebogruppen. For barneavføringen ble prøvene fra 10- og 90 dager etter fødsel valgt og parene med de to avgitte morsmelksprøvene og avføringsprøven fra mor tatt ved 90 dager etter fødsel, slik at det tilhørte fem prøver per mor- barnpar som til sammen ga 100 prøver. Det ble anskaffet 89 ulike prøver med bakterielt DNA fra Genetic Analysis AS (GA) i Oslo (stammesamling). Seleksjonen av DNA fra ulike fylum ble gjort tildels tilfeldig, samtidig som spesifikke bakterier kjent for å kolonisere morsmelk og avføring ble plukket ut. En oversikt over prøvematerialet fra pro-PACT vises i figur 2.1.



Figur 2.1 Skjematisk fremstilling av prøvematerialet fra pro-PACT studiet. Avføringen og morsmelken er isolert fra 278 mødre/barn par fordelt i probiotikagruppen med 138 mødre og placebogruppen med 140 mødre. Fra hvert par er det plukket ut til sammen fem prøver, morsmelk og avføringsprøver tatt ved 10- og 90 dager etter fødsel i tillegg til en avføringsprøve fra mor tatt 90 dager etter fødsel.

2.2 Polymerase kjede reaksjon(PCR)

Reagenser og primere benyttet for PCR og sekvensering finnes henholdsvis i vedlegg A og vedlegg B.

Den opprinnelige protokollen fra Saruta et al (1995) ble modifisert og syklustiden for hybridisering og ekstensjon ble endret. Det genomiske materialet brukt hadde en konsentrasjon på mellom 1-10 ng/μl, hvor det bakterielle DNA fra GA hadde en konsentrasjon på 5 ng/μl. En mikroliter DNA ble blandet med en PCR-master miks bestående av følgende reaksjonsblanding: 1,25 U HOT FIREpol® DNA polymerase, 1x HOT FIREpol® buffer B2, 2,5 mM Magnesium-diklorid (MgCl₂), 0,2 mM dNTP, 0,2 μM forward primer og 0,2 μM reverse primer i DNase/RNase fritt vann til et sluttvolum på 25 μl per reaksjon. For morsmelksprøvene ble et volum på 2 μl brukt. Reaksjonene ble utført med en initiell denatureringstemperatur på 95 °C i 15 minutter. Temperaturen for denaturering og elongering som ble brukt var henholdsvis 95 °C og 72 °C. Optimal hybridiserings temperatur ble satt til 55 °C. Navn på primere, tid for denaturering elongering og hybridisering, samt antall sykluser brukt er gitt i tabell 2.1.

Tabell 2.1 Oversikt over Primere brukt til DNA amplifikasjon og sekvensering.

| Primersett navn | Denaturerings tid, min:sek | Hybridiserings tid, min:sek | Elongerings tid, min:sek | Antall sykluser | Referanser |
|-----------------|----------------------------|-----------------------------|--------------------------|-----------------|-------------------|
| ITS | 1:00 | 1:00 | 1:00 | 30 | Saruta et al.1997 |
| tm | 0:30 | 0:45 | 1:30 | 35 | Kelly et al. 2001 |
| ITS-Illumina | 1:00 | 1:00 | 1:00 | 12 | Denne oppgaven |
| PRK | 0:30 | 0:30 | 1.20 | 30 | Yu et al.2005 |

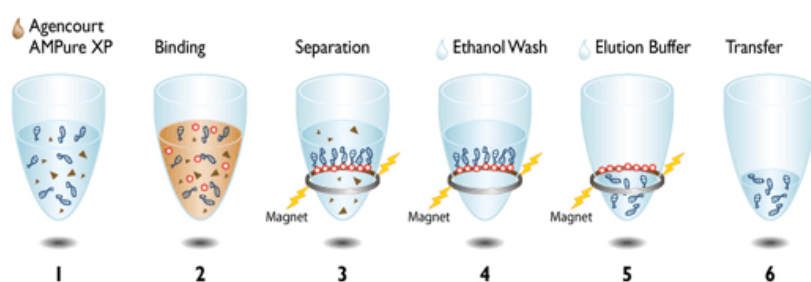
2.2.1 Gradient PCR

For gradient PCR optimaliseringen ble samme reaksjonsblanding og temperaturer brukt som i kapitel 2.2. Temperatur-intervallene ble satt til 45-66 °C ved å sette gradienten til 21 °C og hybridiseringstemperaturen til 55,5 °C. Reaksjonen ble utført med en initiell denaturering ved 95 °C i 15 min og samme tids-spesifikasjoner som i tabell 2.1 for ITS primersettet. Amplikonene ble kontrollert med 1 % agarose gel.

2.2.2 Opprensing av PCR-produkt med AMPure® XP

For å kunne fjerne overflødig primere, primedimere, nukleotider, salter og polymerase ble PCR-amplikonene rensert manuelt med AMPure® XP kit. Denne metoden baseres på SPRI kuler som er paramagnetiske (magnetisk bare i et magnetisk felt), og hver kule er laget av polystyren, omgitt av et lag med magnetitt, som er belagt med karboksyl-molekyler. Det er disse som reversibelt binder DNA i nærvær av polyetylen-glykol (PEG) og salt (vanligvis NaCl). PEG fører til at negativt ladet DNA bindes til karboksylgruppene på kuleoverflaten (Witek et al. 2006). Hvis forholdet endres, endrer dette også på hvilken lengde fragmentene har som binder/ikke binder seg til kulene. Ved lavere forholdet mellom kulene og DNA i løsningen vil det være tilstedet flere lengre fragmenter ved eluering, siden mindre fragmenter holdes igjen i bufferen og kastes i separasjonssteget tidlig i prosessen. En del av årsaken til denne effekten er at DNA fragmentets størrelse påvirker den totale ladningen per molekyl, hvor større DNA molekyler har større ladning og dette fremmer den elektrostatiske interaksjonen med kulene og fortrenger mindre DNA-fragmenter. Det ble tatt ut 15µl av hvert PCR-produkt og overført til en ny mikrobrønnplate. Til hver brønn ble det tilsatt 15µl med

romtemperert AMPure® XP kuler. Prøvene ble pipettert forsiktig opp og ned 10 ganger før de ble inkubert ved romtemperatur i fem minutter for å binde DNA. Etter inkubering ble platen plassert i et magnetisk stativ i 2 minutter for å separere kulene fra løsningen, supernatanten ble pipettert ut uten å berøre de separerte kulene mens platen fortsatt befant seg i det magnetiske stativet. Med platen plassert i det magnetiske stativet ble det tilsatt 100 µl nylaget 70 % etanol til hver brønn av reaksjonsplaten og inkubert i 30 sekunder ved romtemperatur. Etanolen ble så pipettert ut og kastet, vaskeprosessen ble gjentatt to ganger totalt. Da etanolen kunne inneholde rester av kontaminering var det viktig å fjerne all etanolen fra bunnen av brønnene. Reaksjonsplaten ble beholdt i det magnetiske stativet og kulene lufttørket i 15-20 minutter ved romtemperatur til etanolen var fordampet. Platen ble fjernet fra det magnetiske stativet og det ble tilsatt 31,5 µl av elueringsbuffer (dH₂O) til hver brønn av reaksjonsplaten og løsningen ble blandet ved å pipettere den opp og ned 10 ganger før platen ble inkubert ved romtemperatur i 2 minutter. Etter inkubering ble reaksjonplaten plassert i det magnetiske stativet i 2 minutter eller til supernatanten hadde fått en klar farge. Uten å berøre kulene ble 30 µl av supernatanten overført fra hver brønn i reaksjonsplaten til en ny plate. Figur 2.2 viser de ulike stegene ved DNA separasjon med SPRI kuler.



Figur 2.2 SPRI teknologien illustrert med Agencourt AMPure XP. Figuren viser de ulike stegene som utføres ved en separasjon med SPRI kuler. 1. PCR reaksjonen uten tilsatt AMPure XP kuler. 2. Binding av PCR amplikoner til de magnetiske kulene. 3 Separasjon av PCR amplikoner bundet til magnetiske kuler fra kontamineringer. 4. Vask av PCR amplikoner med etanol. 5. Eluering av PCR-amplikoner vekk fra de magnetiske kulene 6. Overføring av rensset PCR produkt til ny plate/rør. Figuren er tatt fra Beckmans offisielle hjemmeside.

2.3 DNA sekvensering

Det ble konstruert et primerbibliotek fra ITS primerparet (ITSF og ITSr) for Illumina. Totalt inneholdt biblioteket 12 reverse primere og 8 forward primere. Det unike Illuminamerket plassert mellom koloniamplifiseringsregionene og Illumina-sekvenserings-primeren på reversprimeren ble konstruert ved å reverskomplimentere indekssekvensene gitt i TruSeq index sequences, indeks 1-12. Indekskoden i forward primeren er gitt i indeks 13- 20 i TruSeq Index Sequences. For å kunne utføre sekvensering med Illumina teknologien måtte alle PCR-produktene amplifisert med ITS-Illumina primersett blandes til en prøve for multipleksing. Konsentrasjonen av DNA varierte for de ulike prøvene, derfor var det nødvendig å utføre en normalisering på bakgrunn av kvantifisering med et spektrofluorometer (Qubit®). En DNA konsentrasjon på 50 ng, 100 ng og 200 ng for henholdsvis avføringsprøvene, kontroller og morsmelkprøvene ble bestemt. Det ble regnet ut volumer for hver prøve som skulle tilfredsstille kravet om de ulike DNA konsentrasjonene. Et utregningseksempel som viser hvilke volumer som ble tilsatt fra hver blanding for å oppnå den bestemte konsentrasjonen finnes i vedlegg F. Det endelige DNA biblioteket ble blandet sammen til en sluttkonsentrasjon på 4 nM etter 16S Metagenomic Sequencing Library preparation protokollen for MiSeq fra Illumina. Amplikonbiblioteket ble applisert på Illumina plattformen MiSeq® og sekvensert med en sluttkonsentrasjon på 10 pM, og MiSeq reagent kit v3 (600-cycle) fra Illumina ble benyttet.

2.4 DNA kvalifisering og kvantifisering

2.4.1 Agarose gelelektroforese, kvalifisering av PCR produkt.

For å bekrefte tilstedeværelsen og lengden på amplikonene etter PCR, ble PCR-produktene separert igjennom en agarosegel ved elektroforese. DNA er negativt ladet ved nøytral pH, da deoksyribosen i DNA molekylet er koblet til en fosfatgruppe. Den negative ladningen fører til at DNA vandrer mot den positive terminalen (anoden) når den befinner seg i et elektrisk felt. Agarosegelen inneholder porer som tillater separasjon av DNA fragmenter basert på størrelse, struktur og ladning. For å visualisere DNA kan et fargestoff benyttes som danner et kompleks med DNA, dette komplekset fluoriserer under bestråling med UV-lys. I denne oppgaven er det brukt en agarosekonsentrasjon på 1 % blandet ut i TAE-buffer. En mikroliter Loading buffer bestående av glyserol, fargestoff og GelRed™ ble tilsatt til 5 µl av hver prøve før

påføring på gel. En 100 bp DNA ladder ble benyttet til å bestemme PCR-produktets størrelse og konsentrasjon basert på sammenligning med størrelsesstandarden (ladder). For visualisering av DNA-GelRed™ komplekset ble gel-dokumentasjons instrumentet Molecular Imager® Gel Doc™ XR Imaging systemet benyttet. En fullstendig protokoll for støping av agarosegel finnes i vedlegg E.

2.4.2 DNA kvantifisering med Qubit

Qubit® fluorometer er et instrument som brukes til kvantifisering av DNA, RNA og protein. Metoden tar i bruk fluoriserende farge for å bestemme konsentrasjonen av nukleinsyrer og protein i en prøve. Ved binding til DNA, sannsynligvis ved interkalering mellom basene, inntar fargen en mer rigid form og blir intenst fluoriserende. Mengden av fluoriserende signal fra blandingen mellom farge og DNA er direkte proporsjonal med konsentrasjonen av DNA i løsningen. Fluorometeret kan plukke opp dette signalet og bestemme DNA konsentrasjonen i prøven ved hjelp av DNA-standarder med kjent konsentrasjon. I denne oppgaven ble dsDNA HS (high sensitivity) analysen brukt, og instrumentet ble kalibrert på nytt før bruk ved hjelp av standardene. En mikroliter fra hver prøve ble kombinert med 199 µl arbeidsløsning jamfør protokollen fra produsenten som vises i vedlegg E.

2.4 Dataanalyse

For å kunne analysere utbytte etter sekvensering og bestemme tilstedeværelsen av bakteriestammer ble det laget en modell for ITS sekvensene tilgjengelig fra genomdatabaser på NCBI. ITS regionen befinner seg i området mellom 16S og 23S området på rRNA genet, derfor ble det med start ifra 16S regionen ekstrahert ut et 300 bp langt området som ble inkludert i modellen fra de tilgjengelige genomene. Det ble laget en modell med såkalte operasjonelle taksonomiske enheter (OTU) ved hjelp av en Prinsipal komponent analyse (PCA). De 300 bp lange sekvensene ble delt opp i heksamerer hvor leserammen ble flyttet en og en nukleotide av gangen. Nedenfor følger det et eksempel for illustrasjon.

Sekvensen `AAATAAACC` kan deles opp i følgene heksamerer: `AAATAA`, `AATAAA`, `ATAAAC`, `TAAACC` og ingen `ACAAAC` for eksempel.

En tabell med disse heksamerene kan representere sekvensen som `1 1 1 1 0`. Det finnes 4^6 mulige kombinasjoner av heksamerer. Disse listes opp i hver sin kolonne, og hver sekvens i en rad. Hver sekvens skannes og radene fylles opp med `1` for hver heksamer som passer jamfør eksempelet illustrert i tabell 2.2. Grunnen til at dette ble gjort er fordi det er mye enklere og beregningsmessig mindre krevende å arbeide med tall enn med tegn. Det er på dette datasettet som genereres at PCA analysen utføres.

Tabell 2.2 Viser hvordan ulike heksamerer scores etter hvor ofte de opptrer i en sekvens.

| | `AAATAA` | `AATAAA` | `ATAAAC` | `TAAACC` | `ACAAAC` |
|--------------|----------|----------|----------|----------|----------|
| `AAATAAACC` | 1 | 1 | 1 | 1 | 0 |
| `AATTACAAAC` | 0 | 0 | 0 | 0 | 1 |

Det er ønskelig å lage en modell som gir informasjon om hvordan de ulike heksamerene passer inn i sekvensene som er grunnlaget for OTU'er ved å gi dem en enkelt score. Dette oppnås ved å utføre en PCA analyse på datasettet for å gi en enkelt poengsum til hver sekvens som er kombinert av de ulike heksamerene. PCA ville komprimere informasjonen i såkalte prinsipal komponenter, som hver for seg har noe innvirkning fra hver faktor som i dette tilfelle er de ulike heksamerer. Kolonner som viser samme mønster komprimeres til en kolonne da disse ikke bidrar til variasjon. Hvor mye hver av faktorene bidrar er definert ved såkalte belastninger (loadings) jamfør seksjon 1.6 i introduksjonen. I analysen av sekvenseringsdata for denne oppgaven ble fem prinsipal komponenter benyttet. Rådataen ble filtrert ved å bruke en Phred kvalitets score > 30 , og ingen mutasjon i barkoden til primeren ble tillatt. Dataene scores og plottes i et BI-plot etter koordinater de får og et punkt utgjør en OTU. Prøvene som analyseres projiseres så på modellen uten å påvirke denne og plasseres i plottet ut ifra hvor de passer best. Dette kan generere færre eller flere OTU'er enn i den opprinnelige modellen. Plasseres en prøve nærme et eksisterende punkt i modellen kan det indikere likhet mellom disse. Avstanden fra en OTU til nærmeste referanse i genomdatabasen er representert med en avstandsscore. OTU'ene blir taksonomisk identifisert ut fra sine nærmeste referansenaboer. Analysen av rådata ble utført i MATLAB.

3. Resultater

3.1 Sammenligning av ITS og tmRNA primersettene

Sammenligningen av kandidat primersettene ble gjort ved å amplifisere sekvensene for 16S rRNA-23S rRNA spacer regionen og tmRNA regionen.

Tabell 3.1 Spesifisiteten av universelle ITS og tmRNA primere.¹⁾

| Bakterieart | tmRNA | ITS |
|-------------------------------|---------------------|-------|
| <i>Escherichia coli</i> | + + - | + + + |
| <i>Bifidobacterium longum</i> | - ²⁾ - - | + + + |

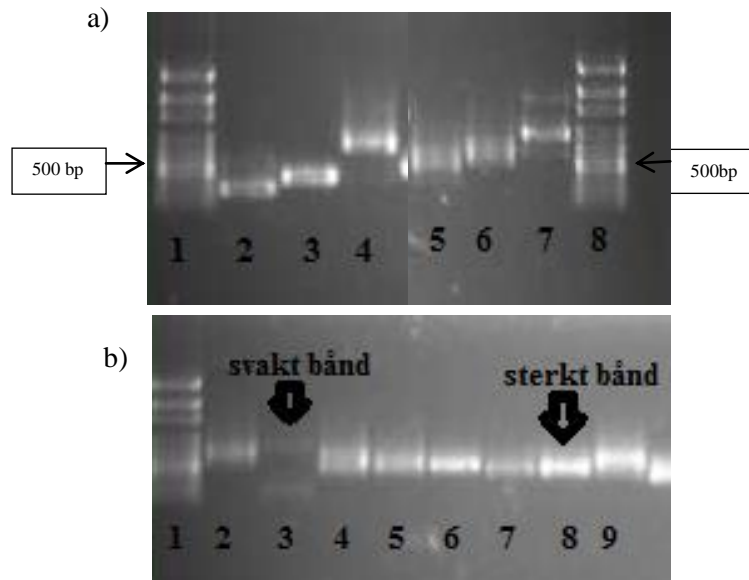
¹⁾ Forsøket ble utført med tre paralleller for hver bakterieart. +, -: amplifisering eller ikke amplifisering hhv og hvert tegn indikerer en parallell

²⁾ Multiple bånd ble detektert.

PCR-produkt amplifisert med ITS primersett ble detektert for både *E. coli* og *B. longum*, mens PCR-produkt amplifisert med tmRNA primersett ble detektert for to paralleller med *E. coli*, men ingen paralleller for *B. longum*, som vist i tabell 3.1. De amplifiserte produktene for begge primersettene varierte i størrelse for begge testorganismene, en illustrasjon finnes i vedlegg D. I videre forsøk ble ikke tmRNA primere inkludert.

3.2 Spesifisitet av ITS primere.

Sensitivitet og spesifisiteten til ITS primersettet ble videre evaluert ved å benytte ulike bakterie arter til amplifisering, 85 av 89 arter ble amplifisert med primersettet, en oversikt over amplifiseringen vises i tabell 3.2. Figur 3.1 a) viser et utklipp som illustrerer lengde diversiteten av ITS regionen amplifisert innad en slekt, det komplette gelbildet vises i vedlegg D. Lengden av de ulike spacerne varierte mellom 350-550 bp og flere PCR produkter kunne observeres for ulike arter. Bakterieartene benyttet for amplifisering med ITS primersettet ble i tillegg brukt til amplifisering med det universale primersettet for prokaryote organismer. Amplifiseringen med PRK primersettet resulterte i en detektering av 87 fra 89 ulike arter. Ingen av de artene som ikke ga amplifisering med PRK korresponderte med arter som ikke ble amplifiserte med ITS primersettet. PCR-produktene ga en fragmentstørrelse på rundt 500bp, og amplifikasjonen var generelt sterkere enn for amplifikasjonen med ITS primerne, figur 3.1 b) illustrerer hva som menes med sterk og svak amplifisering.



Figur 3.1 a-b a) Prøve 2-7: ITS regionen amplifisert innenfor arten *Clostridium*; prøve1 og 8: 100bp molekylvektstandard. b) Bånd styrken viser hva som menes med sterk- og svak amplifisering.

Det universelle primerparet for prokaryote organismer fanget opp 97,7 % av de ulike artene fra stammesamlingen, hvor av de to artene primersettet ikke amplifiserte tilhørte fylum *Firmicutes*. *Lactobacillus salivarius* og *Streptococcus pyogenes* ble ikke amplifisert med PRK primersettet. ITS primersettet ga amplifisering ved agarosegel elektroforese for 95,5 % av artene, hvor av de fire artene som ikke ga amplifisering var *Faecalibacterium prausnitzii* og *Veillonella atypica* tilhørende fylum *Firmicutes*, samt *Hafnia alvei* og *Yersinia pseudotuberculosis* tilhørende fylum *Proteobacteria*. En fullstendig oversikt over stammesamlingen benyttet i denne oppgaven vises i vedlegg C.

Tabell 3.2. Oversikt over ulike bakteriearter amplifisert med ITS og PRK primersettene 1.

| Fylum | Bakterieart | ITS | PRK |
|--|------------------------------|-----|-----|
| Actinobacteria (100/100) ² | <i>B. longum longum</i> | + | + |
| | <i>B. longum infantis</i> | + | + |
| | <i>B. adolescentis</i> | + | + |
| | <i>B. angulatum</i> | + | + |
| | <i>B. breve</i> | + | + |
| | <i>B. catenulatum</i> | + | + |
| | <i>B. dentium</i> | + | + |
| | <i>B. gallicum</i> | + | + |
| | <i>C. acrimucosum</i> | + | + |
| | <i>E. lenta</i> | + | + |
| Bacteroidetes (100/100) ² | <i>M. curtisii</i> | + | + |
| | <i>P. denticolens</i> | + | + |
| | <i>A. putredinis</i> | + | + |
| | <i>B. cellulosilyticus</i> | + | + |
| | <i>B. intestinalis</i> | + | + |
| | <i>B. fragilis</i> | + | + |
| | <i>B. thetaiotaomicron</i> | + | + |
| | <i>B. uniformis</i> | + | + |
| | <i>B. vulgatus</i> | + | + |
| | <i>P. johnsonii</i> | + | + |
| Firmicutes (95.3/95.3) ² | <i>P. marshii</i> | + | + |
| | <i>P. oralis</i> | + | + |
| | <i>B. coecoides</i> | + | + |
| | <i>B. fibrisolvens</i> | + | + |
| | <i>C. asparagiformis</i> | + | + |
| | <i>C. leptum</i> | + | + |
| | <i>C. perfringens</i> | + | + |
| | <i>C. ramosum</i> | + | + |
| | <i>C. saccharolyticum</i> | + | + |
| | <i>D. invisus</i> | + | + |
| Fylum | Bakterieart | ITS | PRK |
| Firmicutes (95.3/95.3) ² | <i>D. longicatena</i> | + | + |
| | <i>E. faecalis</i> | + | + |
| | <i>E. faecium</i> | + | + |
| | <i>E. cylindroides</i> | + | + |
| | <i>E. rectale</i> | + | + |
| | <i>F. prausnitzii</i> | + | + |
| | <i>F. magna</i> | + | + |
| | <i>G. haemolysans</i> | + | + |
| | <i>L. rhamnosus</i> | + | + |
| | <i>L. brevis</i> | + | + |
| Proteobacteria (90.5/100) ² | <i>L. acidophilus</i> | + | + |
| | <i>L. amylolyticus</i> | + | + |
| | <i>L. fermentum</i> | + | + |
| | <i>L. gasseri</i> | + | + |
| | <i>L. paracasei</i> | + | + |
| | <i>L. plantarum</i> | + | + |
| | <i>L. reuteri</i> | + | + |
| | <i>L. salivarius</i> | + | + |
| | <i>L. lactis ssp. Lactis</i> | + | + |
| | <i>L. mesenteroides</i> | + | + |
| Firmicutes (95.3/95.3) ² | <i>L. monocytogenes</i> | + | + |
| | <i>O. valericigenes</i> | + | + |
| | <i>R. indinivorans</i> | + | + |
| | <i>R. albus</i> | + | + |
| | <i>S. aureus</i> | + | + |
| | <i>S. epidermis</i> | + | + |
| | <i>S. infantarius</i> | + | + |
| | <i>S. agalactiae</i> | + | + |
| | <i>S. equinus</i> | + | + |
| | <i>S. mitis</i> | + | + |
| Firmicutes (95.3/95.3) ² | <i>S. oralis</i> | + | + |
| | <i>S. pneumoniae</i> | + | + |
| | <i>S. pyogenes</i> | + | + |
| | <i>V. atypica</i> | + | + |
| | <i>V. dispar</i> | + | + |
| | <i>B. oklahomensis</i> | + | + |
| | <i>C. curvus</i> | + | + |
| | <i>C. yowngae</i> | + | + |
| | <i>D. pigra</i> | + | + |
| | <i>E. aerogenes</i> | + | + |
| Firmicutes (95.3/95.3) ² | <i>E. hormaechei</i> oharad | + | + |
| | <i>E. coli</i> | + | + |
| | <i>H. alvei</i> | + | + |
| | <i>H. pylori</i> | + | + |
| | <i>K. oxytoca</i> | + | + |
| | <i>K. pneumoniae</i> | + | + |
| | <i>N. subflava</i> | + | + |
| | <i>P. mirabilis</i> | + | + |
| | <i>P. vulgaris</i> | + | + |
| | <i>P. aeruginosa</i> | + | + |
| Firmicutes (95.3/95.3) ² | <i>S. bongori</i> | + | + |
| | <i>S. enterica</i> | + | + |
| | <i>S. marcescens</i> | + | + |
| | <i>S. dysenteriae</i> | + | + |
| | <i>S. sonnei</i> | + | + |
| | <i>Y. pseudotuberculosis</i> | + | + |
| | <i>M. orale</i> | + | + |
| | <i>A. muciphila</i> | + | + |
| | <i>F. periodonticum</i> | + | + |
| | Total (95.5/97.7) | + | + |

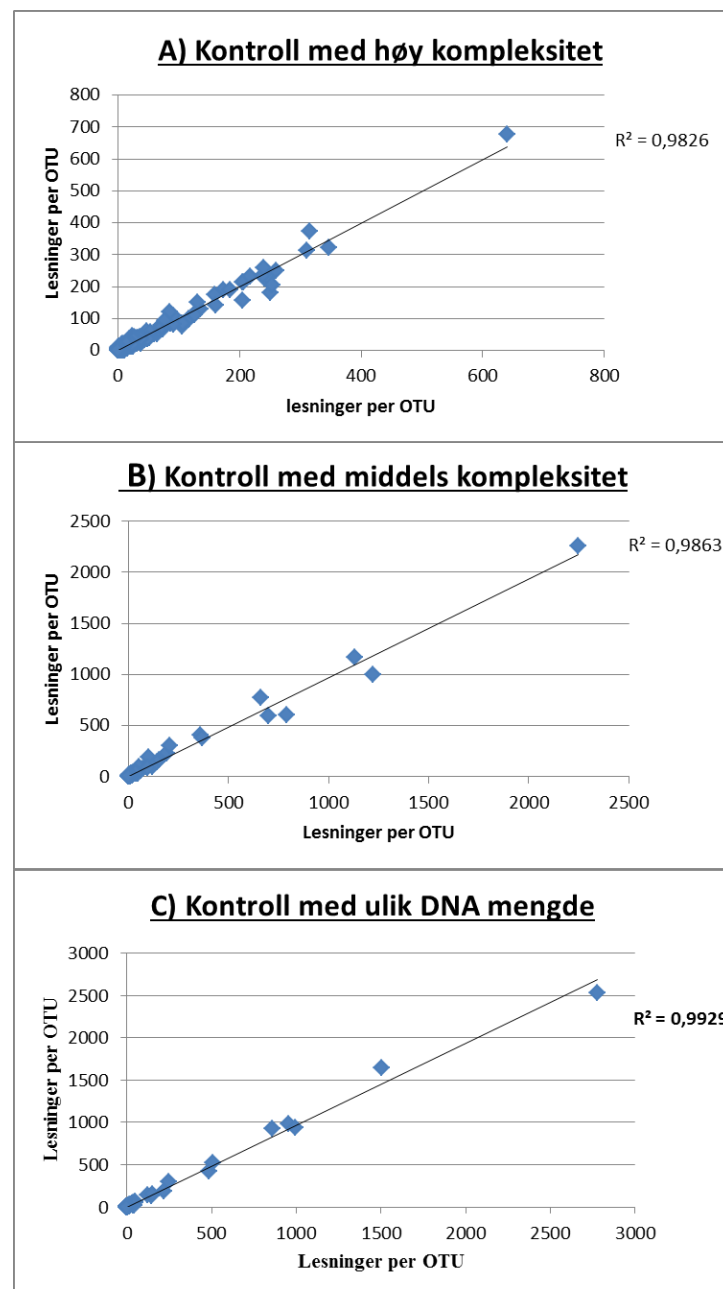
¹ Arten er delt opp og sortert etter fylum.
² ITS, % / PRK, %
³ Firmicutes 100/100
⁴ Firmicutes 100/100
⁵ Firmicutes 100/100

3.3 Illuminasekvensering

3.3.1 Validering av metode

Kontrollene for denne oppgaven bestod av fire ulike sammensetninger, en høy kompleks, en middels kompleks, en med ulikt forhold og en enkelt organisme. Hver kontroll ble amplifisert i to paralleller. Den høy komplekse kontrollen bestod av 83 ulike bakteriearter (figur 3.2 A), mens den middelskomplekse ble satt sammen av 10 ulike bakteriearter (figur 3.2 B). Kontrollen som ble blandet sammen med ulikt forhold i mengden DNA tilsatt, bestod av de samme artene som den middels komplekse, men halvparten av artene ble fortynnet 1:100 fra en utgangskonsentrasjon på 5 ng/μl (figur 3.2 C). Kontrollen som kun bestod av en art ble utgjort av *Ruminococcus albus* og kun en parallell ble tatt med videre i analysen. Det ble utført en regresjonanalyse av kontrollene som bestod av to paralleller og bestemmelseskoeffisienten R^2 var 0,9825, 0,9864 og 0,9929 for henholdsvis kontrollen med høy kompleksitet, kontrollen med middels kompleksitet og kontrollen med ulik DNA mengde.

Det totale antallet OTU'er påvist per kontrollprøve ga et høyere antall enn forventet med hensyn på kompleksiteten kjent. Derfor ble grenseverdien justert til minimum to sekvenser påvist per OTU istedenfor en grenseverdi på en. Denne grenseverdien ga et totalt antall påviste OTU'er som lå nærmere det forventete antallet. I gjennomsnitt ble det funnet 331, 122, 88 og 104 OTU'er ved en grenseverdi på to sekvenser i henholdsvis kontrollen med høy kompleksitet, kontrollen med middels kompleksitet, kontrollen med ulik DNA mengde og kontrollen bestående av en art. Forholdet mellom tilstedeværelsen av OTU'er for begge parallellene og tilstedeværelsen av OTU'er i kun en parallell ble kalkulert for hver kontroll. Forholdene viste seg å være 0,64, 0,61 og 0,65 for henholdsvis kontrollen med høy kompleksitet, kontrollen med middels kompleksitet og kontrollen med ulik DNA mengde. En illustrasjon av forholdene ved ulik grenseverdi vises i vedlegg F.



Figur 3.2 A-C) Regresjonanalyse av kontrollene med grunnlag i antall sekvenser funnet(lesninger) per OTU for parallellene. Svart linje indikerer regresjonslinjen for kontrollene. Den høy komplekse kontrollen bestod av bakteriartene innhentet fra GA med unntak av *Alistipes putredinis*, *Faecalibacterium prausnitzii*, *Hafnia alvei*, *Veillonella atypica*, *Yersinia pseudotuberculosis* og *Leuconostoc mesenteroides*. Kontrollen med middels kompleksitet bestod av artene: *Bacteroides intestinalis*, *Bifidobacterium longum subp infantis*, *Blautia coccoides*, *Clostridium perfringens*, *Escherichia coli*, *Fingoldia magna*, *Lactobacillus acidophilus*, *Lactobacillus reuteri*, *Oscillibacter valericigenes* og *Streptococcus equinus*. Kontrollen med ulik DNA mengde ble satt sammen av de samme artene som i kontrollen med middels kompleksitet, hvor av de fem første nevnt ble fortynnet.

3.3.2 Dominerende OTU'er i prøvematerialet.

Det ble generert 2491 OTU'er totalt i prøvematerialet, der 1356 av disse hadde to eller flere sekvenser. Forekomsten av de ulike OTU'ene ble benyttet til å estimere de ti mest dominerende bakterieartene totalt. De dominerende artene fremstilles i tabell 3.3.

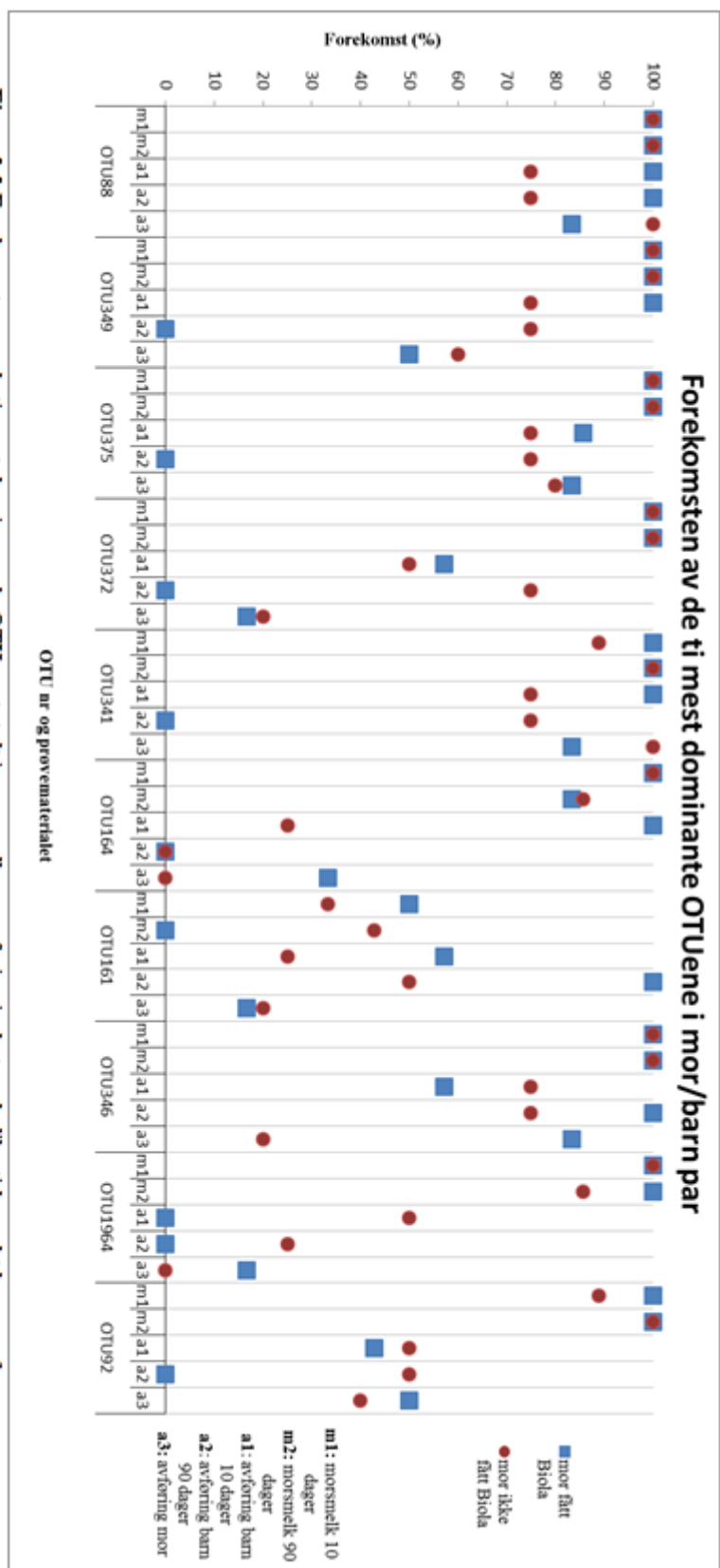
Av de ti mest dominerende OTU'ene totalt bidrar de som adresseres til streptokokker med syv ulike arter som går igjen i flertallet av prøvene. *Streptococcus gordonii* str. *Challis* substr. *CHI* blir funnet igjen i flest prøver totalt. OTU'en for *Marinomonas mediterranea* opptrer hovedsakelig i morsmelk, og OTU'en for *Alkaliphilus oremlandii* opptrer i tilsvarende mengde i både morsmelk og avføring, og begge OTU'ene forekommer i færrest antall prøver. Generelt er det høy forekomst av alle dominerende OTU'er i morsmelk med unntak av OTU 161 (*Alkaliphilus oremlandii*) som forekommer i 40 % og 24 % i henholdsvis prøver isolert 10 dager og 90 dager etter fødsel. I avføringsprøvene fra mødre og barn er forekomsten til de ulike OTU'ene mer varierende enn for morsmelk, hvor av spredningen i avføring fra mødrene skiller seg ut i større grad. Åtte av ti OTU'er ble klassifisert med en scoreverdi på null, mens OTU'en for *Streptococcus gallolyticus* og en av to OTU'er for *Streptococcus dysgalactiae* ssp. *equisimilis* ble tildelt en scoreverdi på en.

Flere av de dominerende OTU'ene ser ut til å vise en sammenheng mellom bakteriearter funnet i morsmelk og i avføring fra barn, som *Streptococcus gordonii* representert av OTU 88. En fremstilling av forekomsten til de ti mest dominante OTU'ene totalt i prøvematerialet innenfor hvert mor- barnpar vises i figur 3.3.

Tabell 3.3 Gjennomsnittlig antall sekvenser og standardavvik innenfor de ti mest dominante bakterie OTUene totalt og deres forekomst i morsmelk og avføring.

| OTU nr | Bakteristamme | Scoring ² verd ³ | Sekvensantall og forekomst ¹ (%) | Morsmelk 10 dager ³ (%) | Morsmelk 90 dager ³ (%) | Avføring 10 dager ³ (%) | Avføring 90 dager ³ (%) | Avføring mor ³ (%) |
|---------|---|--|---|--|---|--|---|--|
| OTU88 | <i>Streptococcus gordonii</i> str. Challis sub str. CH1 | 0 | 936,9 ±121,5 (96,4) | 861,6 ±222,5 (100) 15 ² | 1273,9 ±353 (100) 13 ² | 1923,2 ±579 (90,9) 11 ² | 1352,4 ±604,8 (40) 5 ² | 102,2 ±30,8 (90,9) 11 ² |
| OTU349 | <i>Streptococcus agalactiae</i> 2-22 | 0 | 674,1 ±85,6 (75,3) | 1681,9 ±434,3 (100) | 1072,3 ±297,4 (100) | 73,5 ±22,2 (90,9) | 21,4 ±9,6 (40) | 4,6 ±1,4 (45,4) |
| OTU375 | <i>Streptococcus pneumoniae</i> 70585 | 0 | 666,2 ±84,6 (84,7) | 1677,1 ±433,0 (100) | 981,7 ±272,3 (100) | 102,1 ±30,8 (81,8) | 151,8 ±67,9 (60) | 9,7 ±2,9 (81,8) |
| OTU372 | <i>Streptococcus dysgalactiae</i> s.sp. equisimilis 167 DNA | 1 | 277,6 ±35,3 (64,7) | 223,3 ±57,6 (100) | 1003,7 ±278,4 (100) | 33,5 ±10,1 (54,5) | 40,2 ±18,0 (60) | 6,2 ±1,9 (9,1) |
| OTU341 | <i>Streptococcus agalactiae</i> 2603 V/R | 0 | 266,1 ±33,8 (87) | 162,9 ±42,1 (93,3) | 432,5 ±120,0 (100) | 310,4 ±93,6 (90,9) | 163,4 ±73,1 (60) | 18,2 ±5,5 (90,9) |
| OTU164 | <i>Staphylococcus aureus</i> s.sp. aureus MSHR1132 | 0 | 264,9 ±33,6 (55,1) | 810,8 ±209,3 (100) | 72,8 ±20,2 (84,6) | 260,0 ±78,4 (72,7) | (0) | 2,8 ±0,8 (18,1) |
| OTU161 | <i>Alcaliphilus oremlandii</i> OhILAs | 0 | 245,4 ±31,2 (35,5) | 1,8 ±0,5 (40) | 1,5 ±0,4 (23) | 808,5 ±243,8 (45,4) | 1180,4 ±527,9 (60) | 0,5 ±0,2 (9,1) |
| OTU346 | <i>Streptococcus dysgalactiae</i> s.sp. equisimilis 167 DNA | 0 | 232,8 ±29,6 (83,6) | 529,3 ±136,7 (100) | 449,3 ±124,6 (100) | 26,4 ±7,9 (63,6) | 30,4 ±13,6 (100) | 3,1 ±0,9 (34,5) |
| OTU1964 | <i>Marinomonas mediterranea</i> MMB-1 | 0 | 221,7 ±38,2 (47,9) | 463,4 ±119,6 (100) | 521,6 ±144,7 (92,3) | 0,5 ±0,1 (18) | 0,4 ±0,2 (20) | 0,5 ±0,2 (9,1) |
| OTU92 | <i>Streptococcus gallolyticus</i> s.sp. gallolyticus ATCC 43143 DNA | 1 | 218,6 ±27,8 (72,1) | 120,9 ±31,2 (93,3) | 91,6 ±25,4 (100) | 14,4 ±4,3 (81,8) | 8,6 ±3,8 (40) | 5,5 ±1,6 (45,4) |

¹ Gjennomsnittlig sekvensantall for alle prøver og gjennomsnittets standardavvik, og den prosentvis totale forekomsten av en OTU i prøvematerialet.² Totalt antall prøver i det respektive prøvematerialet.³ Prosentvis forekomst av en OTU per prøvekategori, og gjennomsnittlig sekvensantall per OTU for det respektive prøvematerialet⁴ Scoreverdiene er regnet ut som beskrevet i materialer og metoder.



Figur 3.3 Forekomsten av de ti mest dominerende OTUene totalt i morsmelk og avføring isolert ved ulike tidspunkt hos mor/barn par

14 av 19 mor og barn par hadde både morsmelk og avføringsprøver fra barn med i analysen (figur 3.3). Det kunne påvises tilstedeværelse av samme dominante OTU mellom mor og barn for minst en dominerende OTU hos de 14 parene. De mest stabile OTU'ene med tanke på forekomstens hos både mor og barn var OTU88 (*Streptococcus gordonii* str. *Challis* substr. CH1) og OTU349 (*Streptococcus agalactiae* 2-22), mens OTU161 (*Alkaliphilus oremlandii* OhILAs) og OTU1964 (*Marinomonas mediterranea* MMB-1) viste lavest overlapp mellom mor og barn. Ingen par kunne påvises tilstedeværelsen av alle de dominante OTU'ene hos både mor og barn, og tilstedeværelsen av OTU'ene varierte mellom 2 og 8 per par. I tillegg kunne OTU161 ikke observeres i morsmelksprøver isolert 90 dager etter fødsel. OTU164 (*Staphylococcus aureus*) kunne ikke bli funnet igjen i avføringsprøver isolert fra barn 90 dager etter fødsel. I gjennomsnitt ble det funnet igjen 6 av de 10 dominante OTU'er i både mor og barn for parene der mødrene fikk Biola, og 6,5 av 10 dominante OTU'er i parene der mødrene fikk placebo Biola. Alle OTU'ene adressert til streptokokker ble funnet å være til stede i morsmelkprøvene ved minst et tidspunkt for alle mødrene, og forekomsten så ut til å holde seg stabil over tid for de fleste individene. Forekomsten av de dominante OTU'ene så ut til å endre seg over tid i avføringen fra barn, og avføring fra mor hadde lavest forekomst totalt.

4. Diskusjon

4.1 Evaluering av Illumina ITS sekvenserings metode

Spesifisiteten til ITS primerne ble bekreftet ved utførelse av en PCR med ulike bakteriearter, der iblant arter som vi forventet å finne igjen i morsmelk og avføring. Disse artene ble også amplifisert med det universelle primersettet for prokaryote organismer (PRK) for å kunne utelukke at en eventuell negativ amplifisering skyldtes ITS primersettet. Denne sammenligningen viste høy grad av amplifisering $> 95 \%$ for begge primersettene, med representanter fra alle fylkene som ble testet. Det ble derfor konkludert med at ITS primerne ikke utelukket viktige bakteriegrupper.

Reproduserbarheten av Illumina sekvensresultatene ble gjennomgått ved en regresjonsanalyse utført på kontrollene. Det viste seg at parallellene ga veldig lik informasjon, som ble indikert ved en høy verdi for bestemmelses-koeffisienten (R^2). Det fører til at man kan stole på at en gjentatt utførelse av analysen på samme prøve vil gi likt resultat. Derimot var det totale antallet OTU'er påvist i kontrollene høyere enn forventet. Det er sannsynlig at en overestimering av OTU'er skyldes sekvenseringsfeil. For å minske antallet av falske OTU'er så vi derfor bort fra de med bare en sekvens. Dette forbedret forholdet, men fortsatt ble det påvist over 100 OTU'er for kontrollen som bestod av en bakterieart.

For å unngå falske OTU'er, undersøkte vi fordelingen av de 10 mest dominante OTU'ene i pro-PACT prøvene. Disse var også godt klassifiserte og hadde hovedsakelig en scoreverdi på null. Påvisningen av OTU'er som ble adressert til bakterier som lever i miljøer som ikke har en sammenheng med morsmelk eller avføring gir en indikasjon på feilklassifisering. *Streptococcus pneumoniae* er anerkjent for å forårsake lungebetennelse hos mennesker, men hos friske individer kan den bli funnet igjen i nesehulen uten å skape infeksjoner (Bogaert et al. 2004). Denne bakterie er ikke funnet i morsmelk tidligere, men er nært beslektet med de dominerende streptokokkartene som er funnet i morsmelk, slik at det er nærliggende å tro at den er en feilklassifisert streptokokk (Martín et al. 2004). OTU161 og OTU1964 ble adressert til henholdsvis *Alkaliphilus oremlandii* OhILAs og *Marinomonas mediterranea*, som hadde lavest spredning i pro-PACT datasettet. *A. oremlandii* var tidligere klassifisert som *Clostridium* spp. OhILAs og har blitt isolert fra elv-sedimenter i Ohio (Fisher et al. 2008), mens *M. mediterranea* er en melanin produserende bakterie isolert fra hav (Hernández-Romero et al. 2003). *Clostridium* er funnet igjen i avføring og morsmelk hos henholdsvis

friske barn og mødre, derfor er det nærliggende å tro at *A. oremlandii* er en feilklassifisert *Clostridium* (Collado et al. 2009; Favier et al. 2002; van Nimwegen et al. 2011). Feilklassifisering er en begrensning ved metoden og sannsynligvis skyldes det for lite sekvens informasjon for ITS området. Om 16S rRNA området hadde blitt brukt, ville det vært mye mer sekvensinformasjon tilgjengelig, men begrensningen ligger i at det ikke er mulig å oppnå like god oppløsning. Videre kan feilklassifisering skyldes enkelte forskjeller i ITS sekvensene som ikke ble oppfattet når vi kun valgte ut 300bp lange sekvenser fra referansesgenomene.

Det ble ikke identifisert OTU'er tilhørende fylum *Bacterioidetes*, eller slekten *Bifidobacterium* som er assosiert med både mage-tarm kanalen og morsmelk hos menneske (Albesharat et al. 2011; Collado et al. 2009; Martín et al. 2007). Flere arter tilhørende fylumet *Bacterioidetes* og slekten *Bifidobacterium* var inkludert i stammesamlingen brukt for å bestemme spesifisiteten til ITS primerne. Alle disse ga amplifikasjon som forteller at en ikke identifisering sannsynligvis ikke skyldes primere, eller at prøvene analysert ikke inneholdt fylumet *Bacterioidetes* eller slekten *Bifidobacterium*. Beklageligvis ble ingen arter tilhørende *Bacterioidetes* eller *Bifidobacterium* inkludert i modellen selv om flere av disse var genomsekvenserte. Sannsynligvis skyldtes det at primerregionen inneholdt flere mutasjoner enn det seleksjonskriteriet tillot.

4.2 Påvisning og forekomst av dominerende OTU'er i pro-PACT prøvene

Ingen OTU'er adressert til de probiotiske bakteriene *Lactobacillus rhamnosus* GG, *Lactobacillus acidophilus* La-5 og *Bifidobacterium animalis* subsp *lactis* Bb-12 fra Biola ble funnet igjen blant de dominerende OTU'ene i prøvene. Vi fant at syv OTU'er funnet blant de ti mest dominante, ble adressert til ulike arter innenfor slekten streptokokker. I henhold til litteraturen har streptokokker og stafylokokker ofte blitt funnet igjen som en del av de dominante bakterieartene i morsmelk (Albesharat et al. 2011; Collado et al. 2009; Heikkilä & Saris 2003; Martín et al. 2007; Martín et al. 2004). Hunt et al (2011) kunne påvise at streptokokker og stafylokokker dominerte morsmelk hos de fleste mødre, mens slektene laktobasiller og bifidobakterier ble påvist i mindre grad. Dette resultatet støtter også vår observasjon. Forskjellene funnet i morsmelkens bakterielle sammensetning, kan skyldes genetiske, kulturelle, miljømessige, eller diett forskjeller mellom individene. Videre kan forskjell i lokaliseringen av mødrene og primerne brukt være ansvarlig for de ulike resultatene (Yatsunencko et al. 2012). OTU'ene som ble adressert til streptokokker, kunne observeres i

morsmelkprøvene til alle mødrene ved minst et tidspunkt. Dette funnet kan støtte opp om teorien som hevder at morsmelk kan bestå av en «kjerne» sammensetning av bakterier (Hunt et al. 2011). Det er i kontrast med det som tidligere er påvist for tarmens mikrobiota, der ingen OTU'er i større flertall er delt mellom individer (Turnbaugh et al. 2009). De dominante OTU'ene hadde lavest spredning i avføring fra mor. Dette kan forklares med at konseptet der menneskets tarmflora består av et «kjerne» sett med arter, blir stadig mer usannsynlig, og observasjonen om at mennesker trolig deler en funksjonell «kjerne» mikrobiota virker troverdig (Lozupone et al. 2012). Forekomsten av 8 dominante OTU'er så ut til å være stabile i morsmelken over tre måneder innen en mor. Dette er i tråd med en studie som viste at sammensetningen av bakterier i ulike område av mikrobiotaen i menneske, i et enkelt individ, er godt tilpasset og ofte stabil over tid (Costello et al. 2009).

OTU88 var den mest dominerende OTU'en totalt. Denne OTU'en ble adressert til *Streptococcus gordonii*. *S. gordonii* er en kommensal streptokokk som finnes igjen i munnhulen hos menneske, hvor den kan forårsake plakk (Loo et al. 2000). Det er blitt funnet streptokokker i morsmelk med opprinnelse hos menneskets munnflora i tidligere vurdering av den bakterielle sammensetningen i morsmelk. Disse er også funnet igjen i munnhulen hos spedbarn (Heikkilä & Saris 2003). Dens relative forekomst så ut til å være den samme i morsmelk og avføring hvor mor hadde fått Biola, uavhengig av tidspunkt for isoleringen. Det ble også påvist patogene arter blant de dominerende OTU'ene. Disse kan videreføres fra mor til barn gjennom fødselskanalen, og ved inntak av infisert morsmelk, som kan føre til livstruende infeksjoner hos nyfødte (Bingen et al. 1992; Heikkilä & Saris 2003). *Streptococcus gallolyticus* subsp. *gallolyticus*, *Streptococcus dysgalactiae* og *Staphylococcus aureus* kan påvises i menneskets mage -tarm -kanal og i kjeden hos kvinner (Albesharat et al. 2011; Domelier et al. 2008; Heikkilä & Saris 2003; Tettelin et al. 2002). Litteraturen avslørte at disse tre artene er funnet igjen tidligere i morsmelk, og vår studie kan være med på å bekrefte at de kan forekomme i symptomfrie mødre (Albesharat et al. 2011; Carroll et al. 1979; el-Mohandes et al. 1993; Heikkilä & Saris 2003; Olver et al. 2000). *S. gallolyticus* skilte seg ut med tanke på at dens scoreverdi ble bestemt til å være en. OTU'en for *Streptococcus dysgalactiae* subsp. *equisimilis* fikk tildelt en høy scoreverdi. Sannsynligvis skulle denne klassifiseres som en annen art tilhørende slekten streptokokker, som ikke ble inkludert i modellen

4.2 Overføring fra mor til barn, og bakteriell opprinnelse i morsmelk

Det er kjent at den bakterielle sammensetningen av spedbarnets fekale flora gjenspeiler den bakterielle sammensetningen funnet i morsmelk (Heikkilä & Saris 2003; Martín et al. 2004). I resultatene, som oppsummerte påvisningen av de dominante OTU'ene i de ulike mor -barn par, var det mulig å observere overlapp i mormelk og avføring fra barn (vedlegg G). Da det var mulig å identifisere bakteriene på stammenivå, gir dette en sterk indikasjon på direkte overføring av bakterier mellom mor og barn. Mikrobiotaen hos spedbarn er relativt flyktig og endringer forekommer tidlig i livet, med en økning i mangfold og stabilitet over de tre første leveårene (Favier et al. 2002; Yatsunenko et al. 2012). Denne trenden kunne observeres i forekomsten av OTU'ene i avføringen isolert fra barn, da den så ut til å endre seg med tiden. Det var vanskelig å finne et klart mønster i påvisningen av de dominante OTU'ene med hensyn på om mor hadde fått Biola eller ikke. Det var mulig å se små forskjell på forekomsten i avføringsprøvene, hvor OTU'ene påvises i et færre antall prøver for barn med mødre som ikke fikk Biola. Signifikansen ble ikke regnet ut for disse forskjellene, derfor kan denne trenden ikke bekreftes med sikkerhet. Det har blitt foreslått at kommensale bakterier isolert fra morsmelk, som streptokokker, og probiotiske bakterier har potensialet til å forebygge vekst av *S. aureus* i nyfødte (Heikkilä & Saris 2003; Sprunt & Leidy 1988). Vi observerte at OTU'en adressert til *S. aureus* ble påvist i flere avføringsprøver isolert 10 dager etter fødsel fra barn hvor mødrene hadde fått Biola, enn i prøver hvor mor ikke hadde fått dette tilskuddet. Men OTU'en kunne ikke påvises i avføringsprøvene 90 dager etter fødsel.

Flere teorier eksisterer, om hvordan morsmelken hos menneske koloniseres, og om de nøyaktige mekanismer som fører til at bakteriene når melkekjertlene. Teoriene kan primært deles inn i at påvisningen av bakterier skyldes aktiv migrering eller kontaminering. Aktiv migrerings hypotesen går ut på at i det minste noen av bakterier ifra tarmen hos mor skal kunne nå melkekjertlene gjennom en endogen rute, der dendritiske celler eller makrofager fungere som bærere (Martín et al. 2004). Kontamineringsteorien påstår at morsmelk forurenses med bakterier fra mors hud eller barnets munnhule. En studie har vist at det skjer en høy grad av tilbakegående strømming av morsmelk, tilbake inn i melkekjertlene, under amming (Ramsay et al. 2004). Denne tilbakestrømningen kan påvirke utveksling av bakterier fra barnets munnhule til morsmelk. Det har også blitt vist at den bakterielle sammensetningen i munnhulen hos spedbarn og voksne, er dominert av streptokokker (Aas et al. 2005; Smith et al. 1993). Som nevnt tidligere, var flertallet av dominerende OTU'er i morsmelkprøvene vi

analyserte adressert til slekten streptokokker, hvor av den mest dominante OTU'en totalt (OTU88) stammet fra menneskets munnhule. Selv om denne OTU'en ble påvist i alle avføringsprøver fra mor, er denne observasjonen med på å støtte hypotesen om at bakterier i spedbarn spytt påvirker koloniseringen i morsmelk. Bakterier i munnhulen vil kunne bli overført til mage-tarm-kanalen gjennom spytt og mat, og derfra bli skilt ut gjennom avføring. Marteau et al (2001) så at den bakterielle flora i blindtarmprøver skilte seg kvantitativt og kvalitativt fra den fekal floraen i avføringsprøver. Så selv om avføringsprøvene fra mor kunne påvise OTU88, betyr ikke dette at den er en del av bakteriefloraen i mage-tarm som kan transporteres til melkekjertlene. Det er kjent at bakterier som stammer fra munn, kan entre blodbanen og vandre derfra til hjerte og forårsake endokarditt (Dajani et al. 1997; Sommer et al. 1992). Å hevde at bakterier fra munnhulen til mor kan vandre gjennom blodbanen til melkekjertlene hadde vært en vågal hypotese å hevde, men det finnes ingen studie som kan bekrefte denne påstanden. Flere studier styrker og svekker begge teoriene for kolonisering av morsmelk. Mekanismen bak overføringen av bakterie til melkekjertlene ser ikke ut til å være svart- hvit, og opprinnelsen til i det minste noen av bakteriene funnet i morsmelk kan være arts-avhengige.

5 Konklusjon og videre arbeid

Sekvenseringsteknologien er stadig under utvikling, og bakteriell sekvensinformasjon tilgjengelig øker i takt med denne utviklingen. To primersett ble evaluert med tanke på deres egenskaper som kunne gjøre dem egnet for en stammespesifikk PCR. Bruken av universelle primere, designet for ITS området, virket lovende for å kunne identifisere bakterier på stammenivå. Denne informasjonen vil kunne være med på å utfylle eksisterende sekvensinformasjonen gitt av taksonomiske undersøkelser basert på 16S rRNA genet. Identifikasjon av bakterier på stammenivå var essensielt for å kunne styrke tidligere observasjoner om bakteriell overføring mellom mor og barn. Det vil også kunne være nyttig for debatten om bakteriell opprinnelse i menneskets ulike mikrobiotaer. Når klassifiseringsmodellene har forbedret seg tilstrekkelig, vil det være mulig å undersøke pro-PACT prøvene på nytt for å avdekke tilstedeværelsen av de probiotiske bakteriene fra Biola. Det vil være nødvendig for å kunne si med sikkerhet, at det er bakteriene i Biola som ser ut til å ha en preventiv effekt på allergiutvikling hos barn. Dessverre er det slik at bioinformatiske verktøy ikke klarer å holde følge med den økende sekvensinformasjonen tilgjengelig (Metzker 2010). Derfor vil det kreve mer arbeid for å optimalisere klassifiseringsmodeller, for å sikre en nøyaktig identifisering av bakterier.

Pro-PACT studie bestod av 278 mor -barnpar som fullførte studien, men vi bare så på en liten del av deltagerne. Vi kunne ikke observere karakteristiske trender mellom prøver avhengig av om mor hadde fått Biola eller ikke. Det hadde vært interessant å se på flere deltagere for å kunne fastslå ulikheter i forekomsten av bakterier, mellom Biola og placebo gruppen. Resultatene utvalgt for denne oppgaven reflekterer bare en liten del av datasettet generert. Det ligger mer informasjon der som kan være spennende. Videre arbeid kan være å ta hele datasettet i betraktning og utføre nye analyser, som kan gi et større bilde av ITS områdets evne til å identifisere bakteriestammer. Referansesekvensene som ikke ble fanget opp under OTU genereringen bør inkluderes i modellen. Et gjentak av prøveanalysen etter forbedring av modellen vil kunne gi et resultat som bedre reflekterer bakterieinnholdet i pro-PACT prøvene. I tillegg kan flere mønster for overføring av bakterier mellom mor og barn, samt diversiteten av morsmelk og avføring på stammenivå, avdekkes. For å styrke påstanden om bakteriesammensetningen i morsmelk har sin opprinnelse i bakterier fra munnhulen, vil det være hensiktsmessig å analysere spyttprøver fra mor og barn par.

Referanseliste

- Abdi, H. & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2 (4): 433-459.
- Abrahamsson, T. R., Jakobsson, H. E., Andersson, A. F., Björkstén, B., Engstrand, L. & Jenmalm, M. C. (2012). Low diversity of the gut microbiota in infants with atopic eczema. *Journal of allergy and clinical immunology*, 129 (2): 434-440. e2.
- Albesharat, R., Ehrmann, M. A., Korakli, M., Yazaji, S. & Vogel, R. F. (2011). Phenotypic and genotypic analyses of lactic acid bacteria in local fermented food, breast milk and faeces of mothers and their babies. *Systematic and applied microbiology*, 34 (2): 148-155.
- Andersson, H., Asp, N.-G., Bruce, Å., Roos, S., Wadström, T. & Wold, A. E. (2001). Health effects of probiotics and prebiotics A literature review on human studies. *Food & Nutrition Research*, 45: 58-75.
- Ansorge, W. J. (2009). Next-generation DNA sequencing techniques. *New biotechnology*, 25 (4): 195-203.
- Aso, Y., Akaza, H., Kotake, T., Tsukamoto, T., Imai, K. & Naito, S. (1994). Preventive effect of a *Lactobacillus casei* preparation on the recurrence of superficial bladder cancer in a double-blind trial. The BLP Study Group. *European urology*, 27 (2): 104-109.
- Bennett, R. G., Gorbach, S. L., Goldin, B. R., Chang, T.-W., Laughon, B. E., GREENOUGH, W. B. & Bartlett, J. G. (1996). Treatment of relapsing *Clostridium difficile* diarrhea with *Lactobacillus GG*. *Nutrition Today*, 31 (6): 39S.
- Biasucci, G., Rubini, M., Riboni, S., Morelli, L., Bessi, E. & Retetangos, C. (2010). Mode of delivery affects the bacterial community in the newborn gut. *Early human development*, 86 (1): 13-15.
- Birkeland, S.-E. & Tandberg, A. (2003). Kan probiotiske melkesyrebakterier fremme helsen? *TIDSSKRIFT-NORSKE LAEGEFORENING*, 123 (12): 1701-1703.
- Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A. & Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *science*, 307 (5717): 1915-1920.
- Chuang, L. Y., Cheng, Y. H. & Yang, C. H. (2013). Specific primer design for the polymerase chain reaction. *Biotechnol Lett*, 35 (10): 1541-9.
- Clarridge, J. E., 3rd. (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin Microbiol Rev*, 17 (4): 840-62, table of contents.

- Collado, M., Delgado, S., Maldonado, A. & Rodríguez, J. (2009). Assessment of the bacterial diversity of breast milk of healthy women by quantitative real-time PCR. *Letters in applied microbiology*, 48 (5): 523-528.
- Collins, M. D. & Gibson, G. R. (1999). Probiotics, prebiotics, and synbiotics: approaches for modulating the microbial ecology of the gut. *The American Journal of Clinical Nutrition*, 69 (5): 1052s-1057s.
- Dieffenbach, C., Lowe, T. & Dveksler, G. (1993). General concepts for PCR primer design. *PCR Methods Appl*, 3 (3): S30-S37.
- Dotterud, C. K., Storro, O., Johnsen, R. & Oien, T. (2010). Probiotics in pregnant women to prevent allergic disease: a randomized, double-blind trial. *Br J Dermatol*, 163 (3): 616-23.
- Erlich, H. A. (1989). Polymerase chain reaction. *Journal of clinical immunology*, 9 (6): 437-447.
- Goldin, B. R. (1998). Health benefits of probiotics. *The British journal of nutrition*, 80 (4): S203-7.
- Graffelman, J. & van Eeuwijk, F. (2005). Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research. *Biometrical Journal*, 47 (6): 863-879.
- Gürtler, V. & Stanisich, V. A. (1996). New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. *Microbiology*, 142 (1): 3-16.
- Isolauri, E., Sütas, Y., Kankaanpää, P., Arvilommi, H. & Salminen, S. (2001). Probiotics: effects on immunity. *The American journal of clinical nutrition*, 73 (2): 444s-450s.
- Jolliffe, I. (2005). *Principal component analysis*: Wiley Online Library.
- Kalliomäki, M., Salminen, S., Arvilommi, H., Kero, P., Koskinen, P. & Isolauri, E. (2001). Probiotics in primary prevention of atopic disease: a randomised placebo-controlled trial. *The Lancet*, 357 (9262): 1076-1079.
- Kim, M., Oh, H.-S., Park, S.-C. & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International journal of systematic and evolutionary microbiology*, 64 (Pt 2): 346-351.
- Kostman, J. R., Alden, M. B., Mair, M., Edlind, T. D., LiPuma, J. J. & Stull, T. L. (1995). A universal approach to bacterial molecular epidemiology by polymerase chain reaction ribotyping. *Journal of Infectious Diseases*, 171 (1): 204-208.

- Ku, W., Storer, R. H. & Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30 (1): 179-196.
- Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L. & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences*, 82 (20): 6955-6959.
- Lara-Villoslada, F., Olivares, M., Sierra, S., Miguel Rodríguez, J., Boza, J. & Xaus, J. (2007). Beneficial effects of probiotic bacteria isolated from breast milk. *British Journal of Nutrition*, 98 (S1): S96-S100.
- Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.*, 9: 387-402.
- Martín, R., Heilig, H. G., Zoetendal, E. G., Jiménez, E., Fernández, L., Smidt, H. & Rodríguez, J. M. (2007). Cultivation-independent assessment of the bacterial diversity of breast milk among healthy women. *Research in microbiology*, 158 (1): 31-37.
- McLoughlin, R. M. & Mills, K. H. (2011). Influence of gastrointestinal commensal bacteria on the immune responses that mediate allergy and asthma. *Journal of Allergy and Clinical Immunology*, 127 (5): 1097-1107.
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nat Rev Genet*, 11 (1): 31-46.
- Moore, B. (1981). Principal component analysis in linear systems: Controllability, observability, and model reduction. *Automatic Control, IEEE Transactions on*, 26 (1): 17-32.
- Morey, M., Fernández-Marmiesse, A., Castiñeiras, D., Fraga, J. M., Couce, M. L. & Cocho, J. A. (2013). A glimpse into past, present, and future DNA sequencing. *Molecular Genetics and Metabolism*, 110 (1-2): 3-24.
- Morrow, A. L. & Rangel, J. M. (2004). *Human milk protection against infectious diarrhea: implications for prevention and clinical care*. Seminars in pediatric infectious diseases: Elsevier. 221-228 s.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1992). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *BIOTECHNOLOGY SERIES*: 17-17.
- Noverr, M. C. & Huffnagle, G. B. (2004). Does the microbiota regulate immune responses outside the gut? *Trends in microbiology*, 12 (12): 562-568.

- Ouwehand, A. C., Salminen, S. & Isolauri, E. (2002). Probiotics: an overview of beneficial effects. *Antonie Van Leeuwenhoek*, 82 (1-4): 279-289.
- Penders, J., Gerhold, K., Stobberingh, E. E., Thijs, C., Zimmermann, K., Lau, S. & Hamelmann, E. (2013). Establishment of the intestinal microbiota and its role for atopic dermatitis in early childhood. *Journal of Allergy and Clinical Immunology*, 132 (3): 601-607. e8.
- Rajilić-Stojanović, M., Smidt, H. & De Vos, W. M. (2007). Diversity of the human gastrointestinal tract microbiota revisited. *Environmental microbiology*, 9 (9): 2125-2136.
- Rautava, S., Kalliomäki, M. & Isolauri, E. (2002). Probiotics during pregnancy and breast-feeding might confer immunomodulatory protection against atopic disease in the infant. *Journal of allergy and clinical immunology*, 109 (1): 119-121.
- Saruta, K., Matsunaga, T., Hoshina, S., Kono, M., Kitahara, S., Kanemoto, S., Sakai, O. & Machida, K. (1995). Rapid identification of *Streptococcus pneumoniae* by PCR amplification of ribosomal DNA spacer region. *FEMS microbiology letters*, 132 (1-2): 165-170.
- Shah, N. P. (2001). Functional foods from probiotics and prebiotics: Functional Foods from Probiotics and Prebiotics. *Food technology*, 55 (11): 46-53.
- Shendure, J. & Ji, H. (2008). Next-generation DNA sequencing. *Nature biotechnology*, 26 (10): 1135-1145.
- Smith, L. I. (2002). A tutorial on principal components analysis. *Cornell University, USA*, 51: 52.
- Storro, O., Oien, T., Dotterud, C. K., Jenssen, J. A. & Johnsen, R. (2010). A primary health-care intervention on pre- and postnatal risk factor behavior to prevent childhood allergy. The Prevention of Allergy among Children in Trondheim (PACT) study. *BMC Public Health*, 10: 443.
- Tilsala-Timisjärvi, A. & Alatossava, T. (1997). Development of oligonucleotide primers from the 16S-23S rRNA intergenic sequences for identifying different dairy and probiotic lactic acid bacteria by PCR. *International Journal of Food Microbiology*, 35 (1): 49-56.
- Tsai, C.-C., Lai, C.-H., Yu, B. & Tsen, H.-Y. (2008). Use of specific primers based on the 16S–23S internal transcribed spacer (ITS) region for the screening of *Bifidobacterium adolescentis* in yogurt products and human stool samples. *Anaerobe*, 14 (4): 219-223.

- Vanderhoof, J. A., Whitney, D. B., Antonson, D. L., Hanner, T. L., Lupo, J. V. & Young, R. J. (1999). < i> Lactobacillus GG</i> in the prevention of antibiotic-associated diarrhea in children. *The Journal of pediatrics*, 135 (5): 564-568.
- Vinje, H., Almøy, T., Liland, K. H. & Snipen, L. (2014). A systematic search for discriminating sites in the 16S ribosomal RNA gene. *Microbial informatics and experimentation*, 4 (1): 2.
- Watanabe, S., Narisawa, Y., Arase, S., Okamatsu, H., Ikenaga, T., Tajiri, Y. & Kumemura, M. (2003). Differences in fecal microflora between patients with atopic dermatitis and healthy control subjects. *Journal of allergy and clinical immunology*, 111 (3): 587-591.
- Witek, M. A., Llopis, S. D., Wheatley, A., McCarley, R. L. & Soper, S. A. (2006). Purification and preconcentration of genomic DNA from whole cell lysates using photoactivated polycarbonate (PPC) microfluidic chips. *Nucleic Acids Res*, 34 (10): e74.
- Wood, F. (2009). Principal component analysis.
- Zhang, J., Chiodini, R., Badr, A. & Zhang, G. (2011). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, 38 (3): 95-109.
- Zwieb, C., Wower, I. & Wower, J. (1999). Comparative sequence analysis of tmRNA. *Nucleic acids research*, 27 (10): 2063-2071.

Vedlegg

Vedlegg A: Materialer og utstyr

A.1 Tabell over reagenser og kit laboratorieutstyr og instrumenter. brukt i denne oppgaven.

| Produktnavn | Produsent |
|---|-------------------------------------|
| UltraPure™ Agarose | Invitrogen™ Life Technologies™, USA |
| MilliQ vann | Millipore Corporation, USA |
| Nuklease fritt vann | Amresco®, USA |
| 100 bp ladder | Solis BioDyne, Estland |
| 1 kb ladder | Solis BioDyne, Estland |
| dNTP'er og MgCl ₂ | Solis BioDyne, Estland |
| Alle primere | Invitrogen™ Life Technologies™, USA |
| HOT FIREpol® DNA polymerase | Solis BioDyne, Estland |
| Agencourt®AMPure XP® kit -for PCR rensing | Beckman Coulter, USA |
| GelRed™ | Biotium, USA |
| DNA Loading Dye Buffer Blue 6x | Solis BioDyne, Estland |
| Etanol 96 % | Arcus Kjemi, Norge |
| Qubit™ dsDNA HS Assay kit | Invitrogen™ Life Technologies™, USA |
| PhiX Control v3 | Illumina, USA |
| MiSeq reagent kit v3 (600-cycle) | Illumina, USA |

A.2 Tabell over laboratorieutstyr , instrumenter og programvare brukt i denne oppgaven.

| Navn | Produsent |
|---|-------------------------------------|
| Eppendorfrør, 1,5 ml | Eppendorf, Tyskland |
| Qubit® assay rør | Invitrogen™ Life Technologies™, USA |
| AB-1100 Thermo Fast 96 PCR plate | ThermoScientific, USA |
| Film for 96-brønners plate (AB-0558) | ThermoScientific, USA |
| MiSeq® sequencer | Illumina, USA |
| MasterCycler® gradient | Eppendorf, Tyskland |
| 96 brønns Super Magnet plate | Alpaqua®, USA |
| Molecular Imager® Gel Doc™ XR Imaging System | Bio-Rad laboratories, USA |
| GeneAmp® PCR system 9700 | Applied Biosystems, USA |
| Qubit™ Fluorometer 1.0 | Invitrogen™ Life Technologies™, USA |
| Wide Mini-Sub® Cell GT, gelelektroforese system | Bio-Rad laboratories, USA |
| KuroGEL MaxiPlus 25, gelelektroforesesystem | VWR, USA |
| MATLAB (Matrix Laboratory) | MathWorks |
| Quantity One v.4.6.7 | BioRad |
| Qubit® Data Logger og Qubit® Firmware | Invitrogen™ Life Technologies™, USA |
| Microsoft Excel, Microsoft Word 2010 | Microsoft© |

Vedlegg B: Primer sekvenser

Tabell B.1 Primere benyttet i PCR reaksjoner.

| Navn | Sekvens (5'-3') | Målgangensime | Referanse |
|-----------------|----------------------|---------------|--------------------|
| ITS F | GGTGAAGTCGTAACAAG | Prokaryote | Saruta et al. 1995 |
| ITS R | TGCCAAGGCATCCACC | Prokaryote | Saruta et al. 1996 |
| tm F | GGGGCTGATTCTGGATTCG | Prokaryote | Kelley et al. 2001 |
| tm R | GCTGSGGGAKTTGAACC | Prokaryote | Kelley et al. 2001 |
| PRK341 F | CCTACGGGRBGCASCAG | Prokaryote | Yu et. Al (2004) |
| PRK806 R | GGACTACYVGGGTATCTAAT | Prokaryote | Yu et. Al (2004) |

Tabell B.2 Revers ITSillumina primere, unikt illumina-tag i kursiv.

| nr | Sekvens (5' - 3') |
|----|--|
| 1 | CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 2 | CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 3 | CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 4 | CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 5 | CAAGCAGAAGACGGCATACGAGATCACTCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 6 | CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 7 | CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 8 | CAAGCAGAAGACGGCATACGAGATCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 9 | CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 10 | CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 11 | CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |
| 12 | CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCAAGGCATCCACC |

Tabell B.3 Forward ITS Illumina primere, "primer tag" i kursiv

| nr | Sekvens (5' - 3') |
|----|--|
| 1 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>AGTCA</i> AGGTGAAGTCGTAACAAG |
| 2 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>AGTTC</i> CGGTGAAGTCGTAACAAG |
| 3 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>ATGTC</i> AGGTGAAGTCGTAACAAG |
| 4 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>CCGTC</i> CGGTGAAGTCGTAACAAG |
| 5 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>GTAGA</i> GCGTGAAGTCGTAACAAG |
| 6 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>GTCCG</i> CGGTGAAGTCGTAACAAG |
| 7 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>GTGAA</i> AGGTGAAGTCGTAACAAG |
| 8 | AATGATACGGCGACCACCGAGATCT <u>ACACTCTTTCCCTACACGACGCTCTTCCGATCT</u> <i>GTGGC</i> CGGTGAAGTCGTAACAAG |

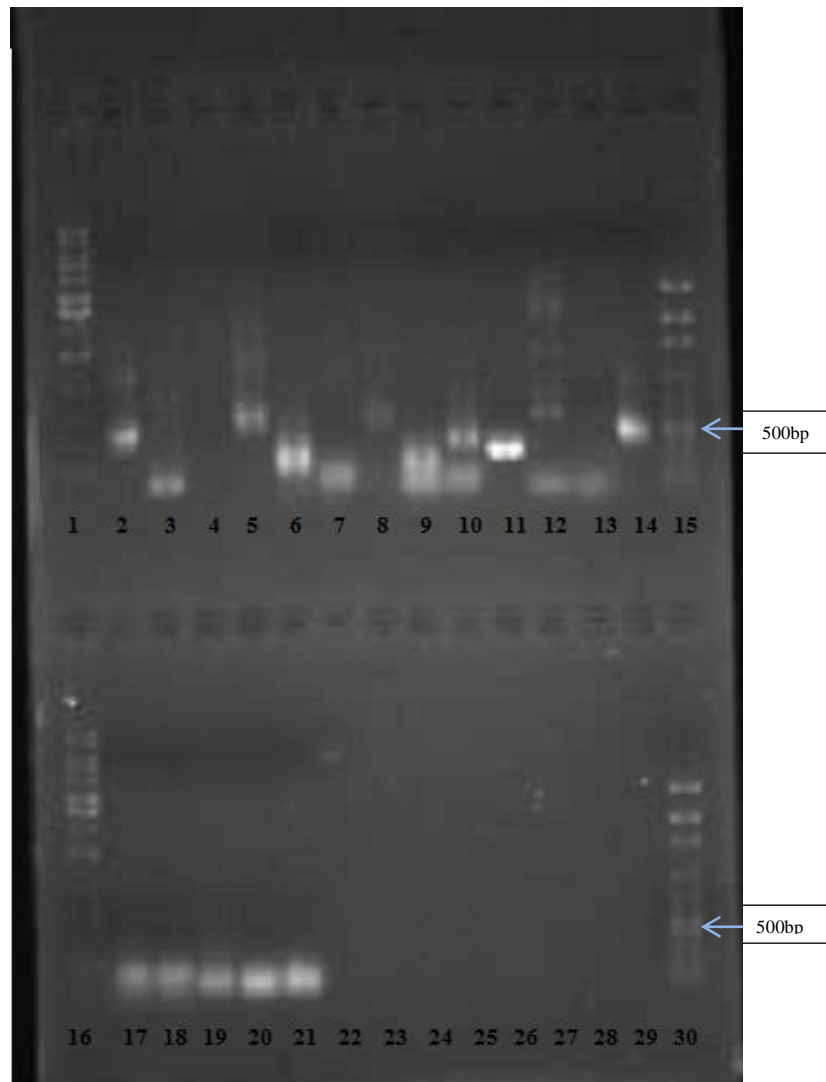
Vedlegg C: Bakteriestammer

Tabell C.1 Stammesamlingen benyttet ved evaluering av ITS primer spesifisitet

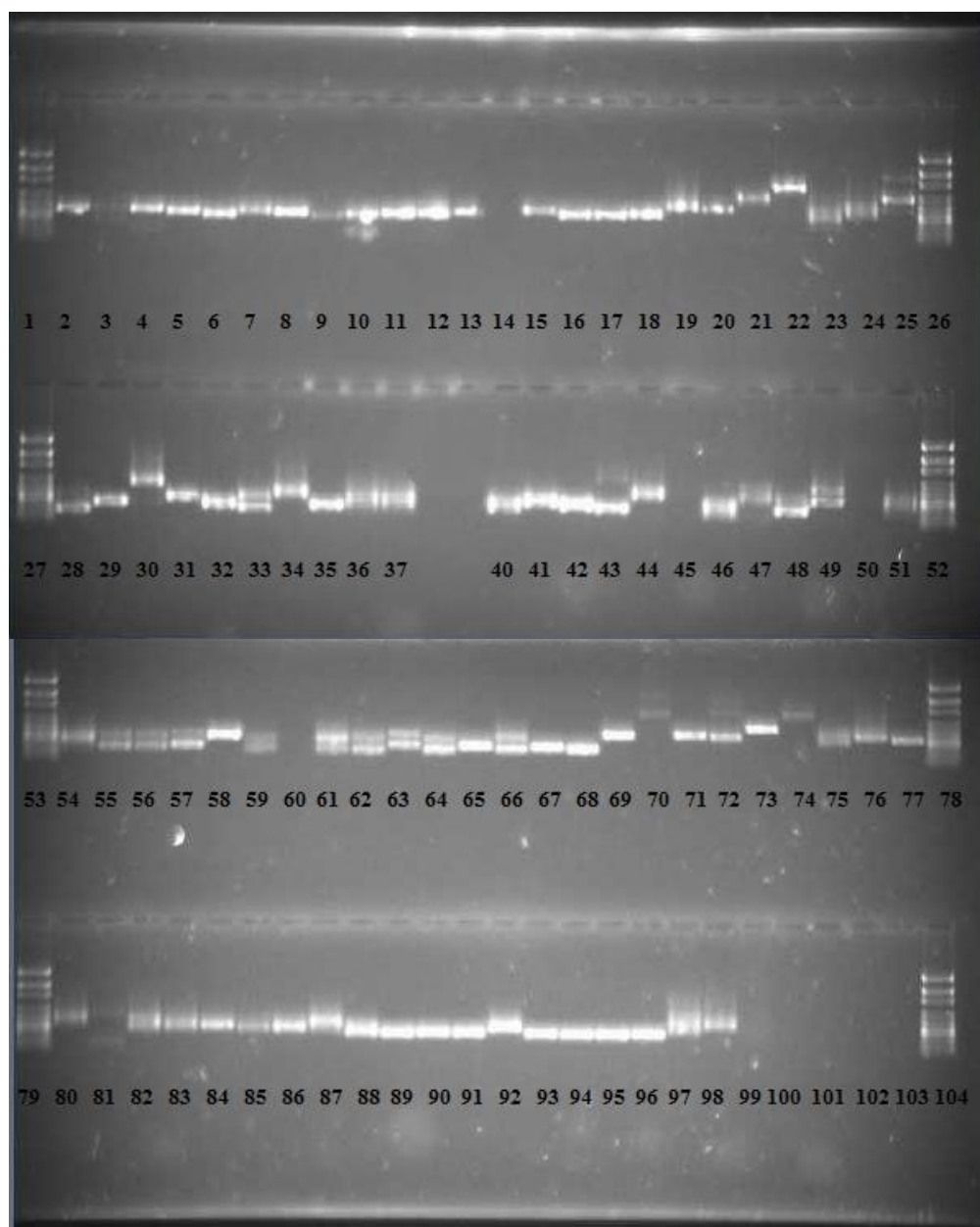
| Art | Prøvenr | Stamme | Referanse |
|---|---------|------------|-------------|
| <i>Akkersmania municipihila</i> | 2 | DSM 22959 | Selma Kreso |
| <i>Alistipes putredinis</i> | 3 | DSM 17216 | Selma Kreso |
| <i>Bacteroides cellulosilyticus</i> | 4 | DSM 14838 | Selma Kreso |
| <i>Bacteroides fragilis</i> | 6 | DSM 2151 | Selma Kreso |
| <i>Bacteroides intestinalis</i> | 5 | DSM 17393 | Selma Kreso |
| <i>Bacteroides thetaiotamicron</i> | 7 | DSM 2079 | Selma Kreso |
| <i>Bacteroides uniformis</i> | 8 | DSM 6597 | Selma Kreso |
| <i>Bacteroides vulgatus</i> | 9 | DSM 1447 | Selma Kreso |
| <i>Bifidbacterium longum subsp infantis</i> | 1 | DSM 20088 | Selma Kreso |
| <i>Bifidbacterium longum subsp longum</i> | 10 | DSM 20219 | Selma Kreso |
| <i>Bifidbacterium adolescentis</i> | 12 | DSM 20083 | Selma Kreso |
| <i>Bifidbacterium angulatum</i> | 13 | DSM 20098 | Selma Kreso |
| <i>Bifidbacterium breve</i> | 15 | DSM 20213 | Selma Kreso |
| <i>Bifidbacterium catenulatum</i> | 16 | DSM 16992 | Selma Kreso |
| <i>Bifidbacterium dentium</i> | 17 | DSM 20436 | Selma Kreso |
| <i>Bifidbacterium gallicum</i> | 18 | CCUG 34979 | Selma Kreso |
| <i>Blautia coccoides</i> | 19 | DSM 935 | Selma Kreso |
| <i>Burkholderia oklahomensis</i> | 20 | DSM 21774 | Selma Kreso |
| <i>Butyrivibrio fibrisolvens</i> | 21 | DSM 3071 | Selma Kreso |
| <i>Campylobacter curvus</i> | 22 | DSM 6644 | Selma Kreso |
| <i>Citrobacter youngae</i> | 23 | DSM 17578 | Selma Kreso |
| <i>Clostridium asparagiforme</i> | 24 | DSM 15981 | Selma Kreso |
| <i>Clostridium leptum</i> | 25 | DSM 753 | Selma Kreso |
| <i>Clostridium perfringens</i> | 28 | DSM 756 | Selma Kreso |
| <i>Clostridium ramosum</i> | 29 | DSM 1402 | Selma Kreso |
| <i>Clostridium saccharolyticum</i> | 30 | DSM 2544 | Selma Kreso |
| <i>Corynebacterium aurimucosum</i> | 31 | DSM 44532 | Selma Kreso |
| <i>Desulfomonas pigra</i> | 32 | DSM 749 | Selma Kreso |
| <i>Dialister invisus</i> | 33 | DSM 15470 | Selma Kreso |
| <i>Dorea longicatena</i> | 34 | DSM 13814 | Selma Kreso |
| <i>Eggerthella lenta</i> | 35 | DSM 2243 | Selma Kreso |
| <i>Enterobacter aerogenes</i> | 35 | DSM 30053 | Selma Kreso |
| <i>Enterobacter hormaechei oharae</i> | 37 | CCUG 53905 | Selma Kreso |
| <i>Enterococcus faecalis</i> | 40 | DSM 20478 | Selma Kreso |
| <i>Enterococcus faecium</i> | 41 | DSM 20477 | Selma Kreso |
| <i>Escherichia coli</i> | 42 | DSM 30083 | Selma Kreso |
| <i>Eubacterium cylindroides</i> | 43 | DSM 3983 | Selma Kreso |
| <i>Eubacterium rectale</i> | 44 | DSM 17629 | Selma Kreso |
| <i>Faecalibacterium prausnitzii</i> | 45 | DSM 17677 | Selma Kreso |
| <i>Fingoldia magna</i> | 46 | DSM 20470 | Selma Kreso |
| <i>Fusobacterium periodonticum</i> | 48 | DSM 19545 | Selma Kreso |
| <i>Gemella haemolysans</i> | 49 | CCUG 53991 | Selma Kreso |
| <i>Hafnia alvei</i> | 50 | DSM 30163 | Selma Kreso |
| <i>Helicobacter pylori</i> | 51 | DSM 21031 | Selma Kreso |
| <i>Klebsiella oxytoca</i> | 54 | DSM 5175 | Selma Kreso |
| <i>Klebsiella pneumoniae</i> | 55 | DSM 30104 | Selma Kreso |
| <i>Lactobacillus acidophilus</i> | 58 | DSM 20079 | Selma Kreso |
| <i>Lactobacillus amylolyticus</i> | 59 | DSM 11664 | Selma Kreso |
| <i>Lactobacillus brevis</i> | 57 | DSM 20054 | Selma Kreso |
| <i>Lactobacillus fermentum</i> | 61 | DSM 20052 | Selma Kreso |

| | | | |
|---|-----|------------|-------------|
| <i>Lactobacillus gasseri</i> | 62 | DSM 20243 | Selma Kreso |
| <i>Lactobacillus paracasei</i> | 63 | DSM 5622 | Selma Kreso |
| <i>Lactobacillus plantarum</i> | 64 | DSM 20205 | Selma Kreso |
| <i>Lactobacillus reuteri</i> | 65 | DSM 20053 | Selma Kreso |
| <i>Lactobacillus rhamnosus</i> | 56 | DSM 20021 | Selma Kreso |
| <i>Lactobacillus salivarius</i> | 66 | DSM 20554 | Selma Kreso |
| <i>Lactococcus lactis subsp. Lactis</i> | 67 | DSM 20481 | Selma Kreso |
| <i>Leuconostoc mesenteroides</i> | 68 | DSM 20346 | Selma Kreso |
| <i>Listeria monocytogenes</i> | 69 | DSM 20600 | Selma Kreso |
| <i>Mobiluncus curtisii</i> | 70 | CCUG 17762 | Selma Kreso |
| <i>Mycoplasma orale</i> | 71 | DSM 25590 | Selma Kreso |
| <i>Neisseria subflava</i> | 72 | DSM 17610 | Selma Kreso |
| <i>Oscillibacter valericigenes</i> | 73 | DSM 18026 | Selma Kreso |
| <i>Parabacteroides johnsonii</i> | 74 | DSM 18315 | Selma Kreso |
| <i>Parascardovia denticolens</i> | 75 | DSM 10105 | Selma Kreso |
| <i>Prevotella marshii</i> | 76 | DSM 16973 | Selma Kreso |
| <i>Prevotella oralis</i> | 77 | DSM 20702 | Selma Kreso |
| <i>Proteus mirabilis</i> | 80 | DSM 30116 | Selma Kreso |
| <i>Proteus vulgaris</i> | 81 | DSM 13387 | Selma Kreso |
| <i>Pseudomonas aeruginosa</i> | 82 | DSM 50071 | Selma Kreso |
| <i>Roseburia inulinivorans</i> | 83 | DSM 16841 | Selma Kreso |
| <i>Ruminococcus albus</i> | 84 | DSM 20455 | Selma Kreso |
| <i>Salmonella bongori</i> | 85 | DSM 13771 | Selma Kreso |
| <i>Salmonella enterica</i> | 86 | DSM 17058 | Selma Kreso |
| <i>Serratia marcescens</i> | 87 | DSM 30121 | Selma Kreso |
| <i>Shigella dysenteriae</i> | 88 | CCUG 9565 | Selma Kreso |
| <i>Shigella sonnei</i> | 89 | DSM 5570 | Selma Kreso |
| <i>Staphylococcus aureus</i> | 90 | DSM 20231 | Selma Kreso |
| <i>Staphylococcus epidermis</i> | 91 | DSM 20044 | Selma Kreso |
| <i>Streptococcus agalactiae</i> | 93 | DSM 2134 | Selma Kreso |
| <i>Streptococcus equinus</i> | 94 | DSM 20558 | Selma Kreso |
| <i>Streptococcus infantarius</i> | 92 | DSM 22957 | Selma Kreso |
| <i>Streptococcus mitis</i> | 95 | DSM 12643 | Selma Kreso |
| <i>Streptococcus oralis</i> | 96 | DSM 20066 | Selma Kreso |
| <i>Streptococcus pneumoniae</i> | 97 | DSM 20566 | Selma Kreso |
| <i>Streptococcus pyogenes</i> | 98 | DSM 20565 | Selma Kreso |
| <i>Veillonella atypica</i> | 100 | DSM 20739 | Selma Kreso |
| <i>Veillonella dispar</i> | 99 | DSM 20735 | Selma Kreso |
| <i>Yersinia pseudotuberculosis</i> | 101 | DSM 8992 | Selma Kreso |

Vedlegg D Agarosegel bilder



Figur D.1 Deteksjon av 16S-23S rRNA spacer region og tmRNA region med universelle primere. Prøve 2-10 ble amplifisert med ITS; prøve 11-14 og 17-21 ble amplifisert med tmRNA primersettet. Prøve 1: 100bp ladder; Prøve 2, 5 og 8: *Escherichia coli*; prøve 3,6 og 9: *Bifidobacterium longum*; prøve 4, 7 og 10: negativ kontroll; prøve 11, 14 og 19: *Escherichia coli*; prøve 12, 17 og 20: *Bifidobacterium longum* ; prøve 13, 18 og 21: negativ kontroll; prøve 15: 100bp ladder ; prøve 16: 1kb ladder; prøve 30: 100bp ladder. Prøve 22-29 er blanke.



Figur D.2 Stammesamlingen amplifisert med ITS primersett for evaluering av spesifisitet. Prøve 1, 26, 27, 52, 53, 78, 79 og 104; 100bp ladder. Prøvenummer i gelbildet korresponderer med prøvenummene vist i tabell C.1.

Vedlegg E Protokoller

E.1 1% agarosegel elektroforese

Reagenser:

50 x TAE- buffer

UltraPure™ Agarose

GelRed™ blandet med 6X Gel Loading Dye (1:2)

100 bp DNA ladder

Fremgangsmåte:

1. 1.0 g agarose ble blandet i 100 ml 50 x TAE-buffer, og varmet i en mikrobølgeovn all agarose var oppløst.
3. En passende støpeform med brønnkam ble festet i en tvinge, og den flytende agaroseløsningen ble helt over. Ved avkjøling stivner løsningen til gel. Etter 20 minutter ble støpeformen med agarosegelen overført til et elektroforesekar med 50 x TAE buffer.
4. I den første brønnen ble størrelsesstandarden bestående av følgende reagenser applisert:
4 µL 100 bp DNA ladder
1 µL GelRed™ blandet med 6X Gel loading dye
5. Prøvene ble applisert til de resterende brønnene; 5 µL PCR-produkt ble blandet med 1 µL GelRed™ og 6X Gel loading dye.
6. Elektroforese ble utført med en spenning på 90 V, til prøvene hadde migrert tilstrekkelig (30-40 min; 90 min for stor gel).
7. DNA fragmentene ble visualisert i et gel-dokumentasjons system.

E.2 Qubit protokoll

Qubit™ Assays

www.invitrogen.com/qubit

NOTE: For best results, store the dye and the buffer at room temperature. Store the DNA, RNA, and protein standards at 4°C. Ensure that all assay reagents are at **room temperature** before you begin.

1. Set up two Assay Tubes for the standards (three for the protein assay) and one tube for each user sample.
2. Prepare the Qubit™ **Working Solution** by diluting the Qubit™ reagent 1:200 in Qubit™ buffer. Prepare 200 µL of **Working Solution** for each standard and sample.
3. Prepare the Assay Tubes* according to the table below.

| | Standard Assay Tubes | User Sample Assay Tubes |
|---|-------------------------|----------------------------|
| Volume of Working Solution (from step 2) to add | 190 µL | 180–199 µL |
| Volume of Standard (from kit) to add | 10 µL | — |
| Volume of User Sample to add | — | 1–20 µL |
| Total Volume in each Assay Tube | 200 µL | 200 µL |

* Use only thin-wall, clear 0.5 mL PCR tubes. Acceptable tubes include Qubit® assay tubes (set of 500, Cat. no. Q32856) or Axygen PCR-05-C tubes (VWR, part no. 10011-830).

4. Vortex all tubes for 2–3 seconds.
5. Incubate the tubes for 2 minutes at room temperature (15 minutes for the Qubit™ protein assay).
6. Insert the tubes in the Qubit® 2.0 Fluorometer and take readings. For detailed instructions, refer to the Qubit® 2.0 Fluorometer manual.
7. *Optional:* Using the Dilution Calculator feature of the Qubit® 2.0 Fluorometer, determine the stock concentration of your original sample.

**QUICK
REFERENCE
CARD**

MP39808 | MAN0002474

Rev. date: 28 Sept 2010

molecular
probes | **invitrogen**

by life technologies™

For technical support, email tech_support@invitrogen.com.
For country-specific contact information, visit www.invitrogen.com.

Figur E.2 Original Qubit protokoll

Vedlegg F Utreknings eksempel for sammensetning av DNA bibliotek og antall genererte OTU'er i kontrollprøvene

Under følger det en utregning som viser hvilke volumer som ble tilsatt fra hver «pool» for å oppnå den ønskede konsentrasjonen av det endelige DNA biblioteket ble blandet sammen for sekvensering med MiSeq fra Illumina:

Lengde på fragmentet(ITS ampliconet): 500bp → gir 1000bp per amplicon

Vekt per basepar: 1bp= 330g/mol

Vekt per amplicon: 1000bp*330g/mol= 330000g bp/ mol

Ønsket sluttkonsentrasjon: $4 \cdot 10^{-9}$ mol/L

Sluttvolum: 1ml= $1 \cdot 10^{-3}$ L

Antall mol: $4 \cdot 10^{-9}$ mol/L * $1 \cdot 10^{-3}$ L= $4 \cdot 10^{-12}$ mol

1 mol har 330000g med basepar.

$4 \cdot 10^{-12}$ mol = $X \div 330000 \text{ g/mol}$ → $X = 4 \cdot 10^{-12} \text{ mol} \cdot 330000 \text{ g/mol}$
 $X = 1,32 \cdot 10^{-6} \text{ g}$ → **1.32 µg**

Andel avføringsprøver: $50 \div 96 = 0,52$ → $(1.32 \text{ µg} \cdot 0,52) \cdot 1000 \text{ ng} = \underline{686,4 \text{ ng}}$

Andel morsmelkprøver: $38 \div 96 = 0,396$ → $(1.32 \text{ µg} \cdot 0,396) \cdot 1000 \text{ ng} = \underline{522,7 \text{ ng}}$

Andel kontrollprøver: $8 \div 96 = 0,083$ → $(1.32 \text{ µg} \cdot 0,083) \cdot 1000 \text{ ng} = \underline{109,6 \text{ ng}}$

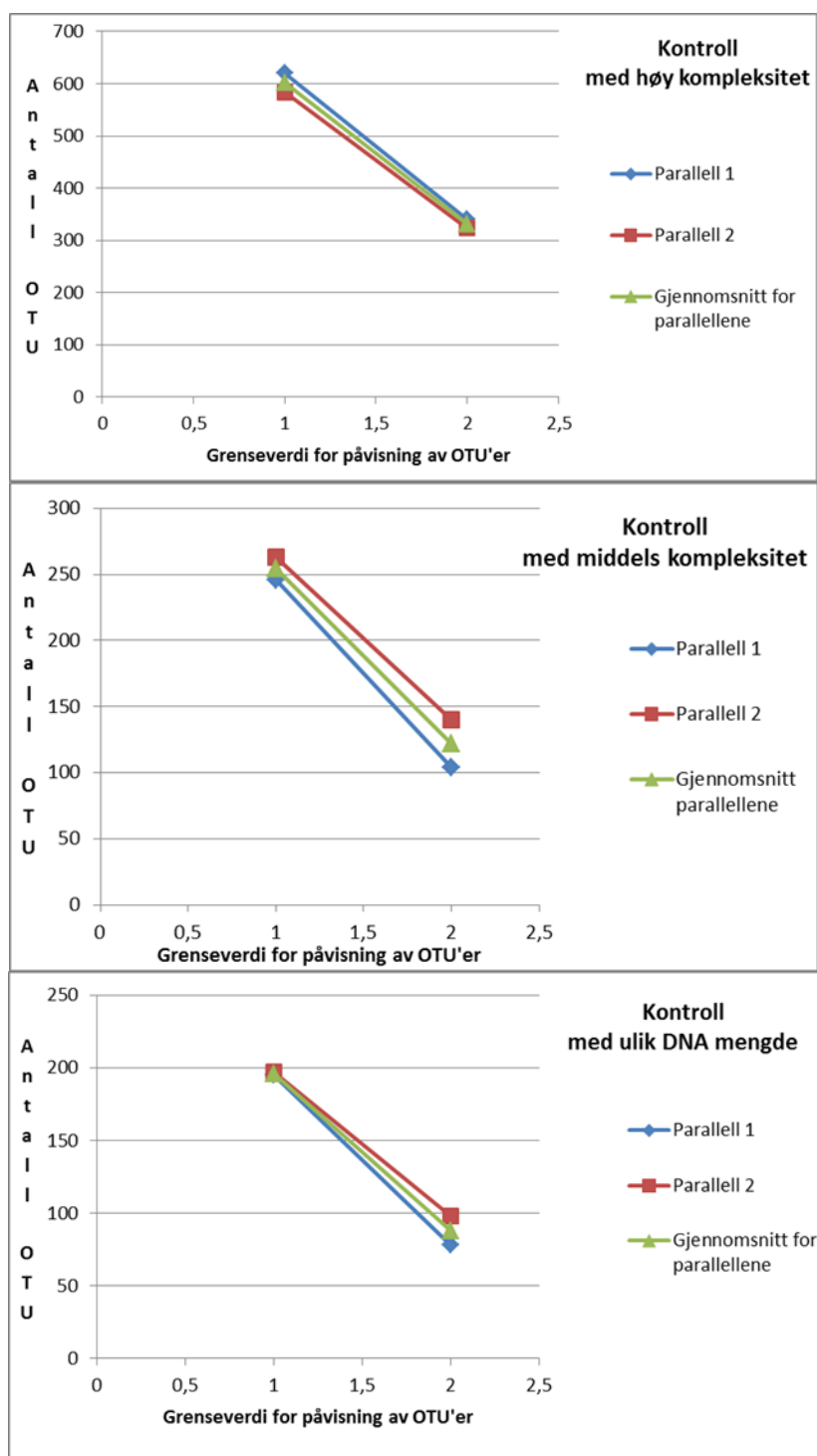
Volum tilsatt fra de ulike prøvene:

Volum avføringsprøver: = $686,4 \text{ ng} \div 50 \text{ ng/µl} = \underline{13,7 \text{ µl}}$

Volum morsmelksprøver: = $522,7 \text{ ng} \div 200 \text{ ng/µl} = \underline{2,6 \text{ µl}}$

Volum kontrollprøver: = $109,6 \text{ ng} \div 100 \text{ ng/µl} = \underline{1,1 \text{ µl}}$

Totalt prøvevolum for en konsentrasjon på 4nM: 17,4 µl DNA+ 982,6µl H₂O



Figur F.1 Antallet OTU'er generert i kontrollene for sekvensering ved to ulike grenseverdier valgt for påvisningen av disse.

Vedlegg G Forekomsten av de ti mest dominante OTU'ene i hvert mor-barnpar

Tabell G.1 Tilstedeværelsen av de ti mest dominante OTU'ene totalt i morsmelk og avføring innenfor ulike mor og barn par¹.

| OTU nr og bakterie-stamme | OTU383 Sordoni str. Chalis str CH1 | OTU349 Sagallactae 2-22 | OTU375 Spneumoniae 70585 | OTU372 S.drygalactae spp.equisimilis 167 DNA | OTU341 Sagallactae 2603V/R | OTU164 S. aureus spp.aureus MISHK132 | OTU161 A.oreumlandii OMLAs | OTU346 S.drygalactae spp.equisimilis 167 DNA | OTU1964 M.mediterranea # MMB-1 | OTU92 S.gallolyticus spp.gallolyticus ATCC 43143 DNA |
|-------------------------------------|--|--|--|--|--|--|--|--|--|--|
| ID nr for mor/barn par ² | | | | | | | | | | |
| 22053 ³ | ++ + 0 ⁴ + | ++ + 0 + | ++ - 0 + | ++ - 0 + | ++ + 0 + | ++ + 0 - | - - - 0 - | ++ 0 - + | ++ - 0 - | ++ - 0 + |
| 20793 | + 0 + + 0 | + 0 + - 0 | + 0 + + 0 | + 0 + + 0 | + 0 + + 0 | + 0 - - 0 | + 0 - - 0 | + 0 + + 0 | + 0 + - 0 | + 0 + + 0 |
| 20850 | + 0 0 + + | + 0 0 + + | + 0 0 + + | + 0 0 + - | - 0 0 + + | + 0 0 + - | - 0 0 - - | + 0 0 + - | + 0 0 + - | - 0 0 + - |
| 21291 | ++ 0 + + | ++ 0 + + | ++ 0 + + | ++ 0 + + | ++ 0 + + | ++ 0 - - | - + 0 + - | ++ 0 + + | ++ 0 + - | ++ 0 - + |
| 21332 | ++ 0 + 0 | ++ 0 + 0 | ++ 0 + 0 | ++ 0 + 0 | ++ 0 + 0 | + - 0 - 0 | - - 0 - 0 | ++ 0 + 0 | ++ 0 - 0 | ++ 0 + 0 |
| 21621 ⁵ | 0 + 0 + 0 | 0 + 0 - 0 | 0 + 0 - 0 | 0 + 0 - 0 | 0 + 0 + 0 | 0 + 0 - 0 | 0 - 0 - 0 | 0 + 0 - 0 | 0 + 0 - 0 | 0 + 0 - 0 |
| 21947 | 0 0 + 0+ | 0 0 + 0 - | 0 0 + 0+ | 0 0 + 0 - | 0 0 + 0 + | 0 0 - 0 - | 0 0 - 0 - | 0 0 + 0 - | 0 0 - 0 - | 0 0 + 0 - |
| 22072 | + 0 + 0 0 | + 0 + 0 0 | + 0 + 0 0 | + 0 - 0 0 | + 0 + 0 0 | + 0 + 0 0 | + 0 - 0 0 | + 0 + 0 0 | + 0 + 0 0 | + 0 - 0 0 |
| 22095 ⁶ | + 0 + 0 + | + 0 + 0 - | + 0 + 0 + | + 0 + 0 - | + 0 + 0 + | + 0 + 0 + | + 0 + 0 + | + 0 + 0 + | + 0 - 0 + | + 0 - 0 + |
| 22199 ⁷ | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | + - 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 |
| 22490 ⁸ | + 0 + + 0 | + 0 + - 0 | + 0 + - 0 | + 0 + - 0 | + 0 + - 0 | + 0 + - 0 | + 0 + + 0 | + 0 - + 0 | + 0 - - 0 | + 0 - - 0 |
| 22508 | ++ 0 0 + | ++ 0 0 + | ++ 0 0 + | ++ 0 0 - | ++ 0 0 + | ++ 0 0 - | ++ 0 0 - | ++ 0 0 - | ++ 0 0 - | ++ 0 0 + |
| 22596 ⁹ | 0 + 0 + - | 0 + 0 + + | 0 + 0 + + | 0 + 0 + - | 0 + 0 + + | 0 - 0 + - | 0 - 0 + - | 0 + 0 + + | 0 + 0 - - | 0 + 0 + - |
| 22670 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 | ++ 0 0 0 |
| 22925 | ++ 0 0 + | ++ 0 0 - | ++ 0 0 - | ++ 0 0 - | ++ 0 0 + | ++ 0 0 - | - - 0 0 - | ++ 0 0 - | + - 0 0 - | ++ 0 0 - |
| 22973 | ++ - - 0 | ++ - + 0 | ++ - - 0 | ++ - - 0 | ++ - - 0 | ++ - - 0 | - + + + 0 | ++ - - 0 | ++ - - 0 | ++ - - 0 |
| 23108 ¹⁰ | 0 + 0 + + | 0 + 0 + - | 0 + 0 + + | 0 + 0 - - | 0 + 0 + - | 0 + 0 + + | 0 - 0 - - | 0 + 0 - + | 0 + 0 - - | 0 + 0 + - |
| 23545 ¹¹ | ++ + 0 + | ++ + 0 + | ++ + 0 + | ++ + 0 - | ++ + 0 + | ++ + 0 - | - - + 0 - | ++ + 0 + | ++ + 0 - | ++ + 0 + |
| 23650 ¹² | + 0 + 0 0 | + 0 + 0 0 | + 0 + 0 0 | + 0 + 0 0 | + 0 + 0 0 | + 0 + 0 0 | - 0 - 0 0 | + 0 + 0 0 | + 0 - 0 0 | + 0 - 0 0 |
| Prøvemateriale | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ | mm ³ a ³ a ³ a ³ |

¹ Tilstedeværelsen av en OTU indikeres av +, - indikerer at OTU'en ikke ble funnet.

² Morsmelk isolert 10 dager etter fødsel

³ Morsmelk isolert 90 dager etter fødsel

⁴ Avføring isolert fra barn 10 dager etter fødsel

⁵ Avføring isolert fra barn 90 dager etter fødsel

⁶ Avføring isolert fra mor 90 dager etter fødsel

⁷ Prøve ikke inkludert

⁸ Identifikasjonsnr tildeelt til hvert mor/barn par, + indikerer at mor har fått biolo



Norges miljø- og
biovitenskapelige
universitet

Postboks 5003
NO-1432 Ås
67 23 00 00
www.nmbu.no